

Order Statistics in Digital Image Processing

IOANNIS PITAS AND ANASTASIOS N. VENETSANOPOULOS, FELLOW, IEEE

In recent years significant advances have been made in the development of nonlinear image processing techniques. Such techniques are used in digital image filtering, image enhancement, and edge detection. One of the most important families of nonlinear image filters is based on order statistics. The widely used median filter is the best known filter of this family. Nonlinear filters based on order statistics have excellent robustness properties in the presence of impulsive noise. They tend to preserve edge information, which is very important to human perception. Their computation is relatively easy and fast compared with some linear filters. All these features make them very popular in the image processing community. Their theoretical analysis is relatively difficult compared with that of the linear filters. However, several new tools have been developed in recent years that make this analysis easier. In this review paper an analysis of their properties as well as their algorithmic computation will be presented.

I. INTRODUCTION

Linear processing techniques are very important tools that are used extensively in digital signal/image processing. Their mathematical simplicity and the existence of a unifying linear systems theory make their design and implementation easy. Moreover, linear processing techniques offer satisfactory performance for a variety of applications. However, many digital image processing problems cannot be efficiently solved by using linear techniques.

An example where linear digital image processing techniques fail is the case of non-Gaussian and/or signal-dependent noise filtering (e.g. impulsive noise filtering). Such types of noise appear in a multitude of digital image processing applications. Impulsive noise is frequently encountered in digital image transmission as a consequence of man-made noise sources or decoding errors. Signal-dependent noise is the photoelectron noise of photosensing devices and the film-grain noise of photographic films [1]. Speckle noise that appears in ultrasonic imaging and in laser imaging is multiplicative noise; i.e. it is signal-dependent noise. Another example where linear techniques fail is the case of nonlinear image degradations. Such degradations occur during image formation and during image transmission through nonlinear channels [1], [5]. The

Manuscript received August 29, 1990; revised July 22, 1992.

I. Pitas is with the Department of Electrical Engineering, University of Thessaloniki, Thessaloniki 54006, Greece.

A. N. Venetsanopoulos is with the Department of Electrical Engineering, University of Toronto, Toronto M5S 1A4, Canada.

IEEE Log Number 9206277.

human visual perception mechanism has been shown to have nonlinear characteristics as well [2], [5].

Linear filters, which were originally used in image filtering applications, cannot cope with the nonlinearities of the image formation model and cannot take into account the nonlinearities of human vision. Furthermore, human vision is very sensitive to high-frequency information. Image edges and image details (e.g. corners and lines) have high-frequency content and carry very important information for visual perception. Filters having good edge and image detail preservation properties are highly suitable for digital image filtering. Most of the classical linear digital image filters have low-pass characteristics [3]. They tend to blur edges and to destroy lines, edges, and other fine image details. These reasons have led researchers to the use of nonlinear filtering techniques.

Nonlinear techniques emerged very early in digital image processing. However, the bulk of related research has been presented in the past decade. This research area has had a dynamic development. This is indicated by the amount of research presently published and the popularity and widespread use of nonlinear digital processing in a variety of applications. Most of the currently available image processing software packages include nonlinear techniques (e.g. median filters and morphological filters). A multiplicity of nonlinear digital image processing techniques have appeared in the literature. The following classes of nonlinear digital image/signal processing techniques can be identified at present: 1) order statistic filters 2) homomorphic filters, 3) polynomial filters, 4) mathematical morphology, 4) neural networks, and 5) nonlinear image restoration. One of the main limitations of nonlinear techniques at present is the lack of a unifying theory that can encompass all existing nonlinear filter classes. Each class of nonlinear processing techniques possesses its own mathematical tools that can provide reasonably good analysis of its performance. Cross-fertilization of these classes has been shown to be promising. For example, mathematical morphology and order statistic filters have been efficiently integrated in one class, although they come from completely different origins.

In the following, we shall focus on the description of the order statistics techniques. Although such techniques have been applied to digital signal processing as well, most

of the reported work has been applied to digital image processing. We shall focus our presentation on digital image processing applications, in order to render it more concise. We shall also give links to other nonlinear filter classes, whenever applicable. The class of filters based on order statistics is very rich. The best known filter is the *median filter*. It originates from robust estimation theory. It was suggested by Tukey for time series analysis [6]. Later, it became popular in digital image processing because of its computational simplicity and its good performance. Its statistical and deterministic properties have been studied thoroughly.

Since its first use, several modifications and extensions of the median filter have been proposed. Many of them have solid theoretical foundations from the theory of robust statistics [7]–[9]. However, there are also filters based on order statistics that have ad hoc structures, due to the lack of a powerful unifying theory in the area of nonlinear filtering. Several efforts have been made in the past decade to provide a unifying theory in the area of order statistics filtering. Some fruitful results based on *threshold decomposition* (to be described later) are expected to provide useful design and implementation tools. In general, filters based on order statistics have good behavior in the presence of additive white noise or impulsive noise, if they are designed properly. Many of them have good edge preservation properties.

The adaptation of order statistics filters is a very important task. It is well known that image characteristics (e.g. local statistics) change from one image region to the other. Noise characteristics usually vary with time. Thus, digital image filters based on order statistics must be spatially and/or temporally adaptive. Furthermore, the characteristics of the human visual system (e.g. edge preservation requirements, local contrast enhancement) lead to spatially adaptive digital image filter structures as well. Another reason for the adaptation of the order statistics filters has to do with the difficulties encountered in the optimal design of such filters for certain characteristics of signal and noise. Although order statistics filters are based on rich mathematical foundations, such design algorithms do not exist or are difficult to implement.

One of the main reasons for the popularity and widespread use of certain filters based on order statistics is their computational simplicity. Their computation can become faster if appropriate fast algorithms are designed. Several such algorithms have appeared in the literature, especially for the fast (serial or parallel) implementation of the median filter. Another research activity is the design of special VLSI chips for order statistics filtering. A number of chips for fast median and max/min filtering have been presented in the literature. The related efforts for fast filter implementation are reviewed in this paper as well.

The organization of the paper is as follows. Section II includes the probabilistic and the deterministic properties of the median filter and its modifications. The analysis is rather detailed, since it provides the tools and methods for the evaluation of the performance of the rest of the order

statistics filters. Section III gives an overview of several filters that are based on order statistics. It also includes the description of certain filters that are not based on order statistics but are related to them through robust estimation theory. Adaptive order statistics filters are presented in Section IV. Algorithms and image processor architectures suitable for fast order statistics filtering are given in Section V. Conclusions are drawn in Section VI.

II. THE MEDIAN FILTER AND ITS EXTENSIONS

The median filter is the most popular example of filters based on *order statistics* [4], [5]. Order statistics have played an important role in statistical data analysis and especially in the robust analysis of data contaminated with outlying observations, called outliers [4], [7], [8]. One of the most important applications of order statistics is *robust estimation* of parameters [7]–[9]. The median is a prominent example of robust estimators. In the following, a brief introduction to some concepts of robust statistics will be given. This description aims at giving some definitions that will be used in the description of the robustness properties of the median filter. Let us suppose that the random variables X_i , $i = 1, \dots, n$, are distributed according to a known parametric probability function $F(\mathbf{X}, \Theta)$. Such parameters can be the mean and the standard deviation; i.e., $\Theta^T = [\mu \ \sigma]$. The parameter vector Θ has to be estimated from the observation data X_i , $i = 1, \dots, n$. Mean and standard deviation estimations correspond to *location* and *scale* estimation. The estimators have the form

$$T_n = T(X_1, \dots, X_n), \quad (1)$$

where $T(X_1, \dots, X_n)$ denotes a function of X_1, \dots, X_n . Often, there are some observation data (outliers) that have very different probability distributions from the model $F(\mathbf{X}, \Theta)$. Therefore, the probability distribution model is only approximately valid. Robust estimation theory proposes estimators for approximate parametric models, which minimize the effect of the outliers. An estimator can have two robustness measures. The *breakdown point* ϵ^* ($0 \leq \epsilon \leq 1$) determines the percentage of the outliers, above which the estimator becomes unreliable [7], [8]. If an estimator has breakdown point $\epsilon^* = 1/2$ (e.g. the median), it is reliable only if less than 50% of the observation data outlier from the model distribution. A formal mathematical definition of the breakdown point can be found in [7]. The *influence function* $IF(x; T, F)$ of an estimator shows the effect of a single outlier on the performance of the estimator at a point x . It is defined as follows [7]:

$$IF(x; T, F) = \lim_{t \rightarrow 0} \frac{T((1-t)F + t\Delta_x) - T(F)}{t}. \quad (2)$$

The *gross-error sensitivity* γ^* measures the worst effect of a contamination at any point x :

$$\gamma^* = \sup_x |IF(x; T, F)| \quad (3)$$

with sup denoting supremum. If γ^* is finite, the estimator T is said to be B-robust at the distribution F . The performance of an estimator is directly related to its output

variance, $V(T, F)$. The smaller the variance, the better the performance of the estimator. The relative efficiency of two estimators T, S at data distribution F is measured by their *asymptotic relative efficiency*, $ARE(T, S)$:

$$ARE(T, S) = \frac{V(S, F)}{V(T, F)}. \quad (4)$$

Here $V(S, F)$, $V(T, F)$ are the variances of the estimators T, S at the distribution F . The higher $ARE(T, S)$ is, the better the performance of the estimator T over the estimator S . The calculation of the ARE of various estimators and its relation to the influence function can be found in [7]. The previous description of robust estimation concepts is far from complete. The interested reader can find more details in [5]–[9]. Robust estimation has found significant applications in digital signal processing [14]–[24]. An excellent review in this area is [19]. Having defined some basic notions of robust estimation, we proceed to the definition and analysis of the median filters.

Let X_1, X_2, \dots, X_n be random variables. If they are arranged in ascending order of magnitude,

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}, \quad (5)$$

$X_{(i)}$ is called the i th-order statistic. The maximum and the minimum of X_i , $i = 1, \dots, n$ are denoted by $X_{(n)}$, $X_{(1)}$ respectively. A very important order statistic is the *median* $med(X_i)$, given by

$$med(X_i) = \begin{cases} X_{(\nu+1)} & \text{if } n = 2\nu + 1 \\ (X_{(\nu)} + X_{(\nu+1)})/2 & \text{if } n = 2\nu \end{cases}. \quad (6)$$

A *one-dimensional median filter* of size n , $n = 2\nu + 1$, is defined by the following relation [25]:

$$y_i = med(x_{i-\nu}, \dots, x_i, \dots, x_{i+\nu}), \quad i \in \mathbf{Z}, \quad (7)$$

where \mathbf{Z} denotes the set of integers. In most cases, filter windows having odd length are used, due to the simplicity of the corresponding definition, as can be seen in (6). The signal sequence is of finite extent in most practical cases. It is appended by appropriately chosen values (usually 0's) at both ends in order to accommodate boundary problems. An example of filtering a three-level sequence by a median filter (7) of length $n = 3$ is shown in Fig. 1(a). The input sequence has been appended by one sample at the beginning and at the end, respectively, to create the necessary boundary conditions. A *two-dimensional median filter* has the following definition:

$$y_{ij} = med(x_{i+r, j+s}; (r, s) \in A), \quad (i, j) \in \mathbf{Z}^2. \quad (8)$$

The set $A \subset \mathbf{Z}^2$ is the *filter window*. Such windows are shown in Fig. 1(b). The window shape influences certain spatial properties of the median filter, e.g. edge and image detail preservation. Analog versions of the median filter have also been defined. Their definition and properties can be found in [26] and [27]. Having defined the median filter, we can proceed to the presentation of its properties.

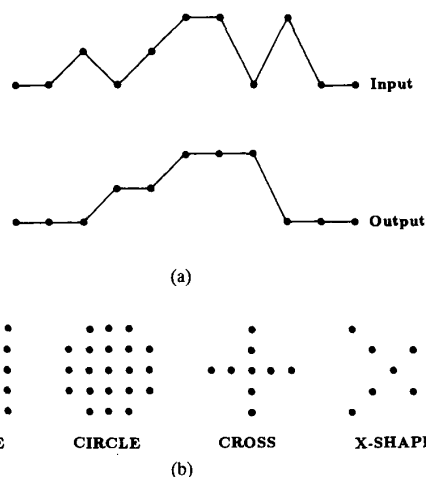


Fig. 1. (a) Filtering of a three-valued sequence by a median filter of length $n = 3$. (b) Windows of two-dimensional median filters.

Statistical analysis presents the statistical properties of the median filter output. Deterministic analysis presents certain structural properties (e.g. the shape of signals that are invariant under median filtering).

A. Statistical Analysis

If X_i , $i = 1, \dots, n$ are independent identically distributed (iid) random variables having cumulative density function (cdf) $P(x)$, the cdf of the r -th order statistic, $X_{(r)}$, is given by

$$F_r(x) = Pr\{X_{(r)} \leq x\} = \sum_{i=r}^n \binom{n}{i} P^i(x) [1 - P(x)]^{n-i}. \quad (9)$$

The probability density function (pdf) $f_r(x)$ of $X_{(r)}$ is given by

$$f_r(x) = n \binom{n-1}{r-1} P^{r-1}(x) [1 - P(x)]^{n-r} p(x), \quad (10)$$

where $p(x)$ is the pdf of X_i , $i = 1, \dots, n$. The cdf and pdf of the median can be found from (9) and (10) by substituting $r = \nu + 1$. It can be easily seen from these relations that the analytic calculation of the cdf and the pdf of the median filter output is cumbersome, even for the iid case. This fact makes the theoretical analysis relatively complicated. Numerical methods for the calculation of the pdf and cdf can be found in [10] and [11]. Order statistic distributions have been used in optimal data sorting [12]. The interested reader is referred to [4] for the calculation of the expected values and moments of order statistics [13]. As stated previously, the output variance is of particular importance to median filter performance. The smaller the variance, the better the filter performance. It has been shown that median filters perform well for *long-tailed* noise distributions (e.g. Laplacian noise), whereas their performance is poor for *short-tailed* noise distributions (e.g. uniform noise) [7]–[9]. This fact suggests that the median filter is efficient at removing impulsive noise. The good performance of the

Table 1 Asymptotic Relative Efficiency of the Median Filter with Respect to the Arithmetic Mean

	Uniform Distribution	Gaussian Distribution	Laplacian Distribution
ARE	0.33	0.637	2.0

median filter at long-tailed distributions is explained by the fact that it minimizes the L_1 norm [5], [9]:

$$\sum_{i=1}^n |x_i - T_n| \rightarrow \min. \quad (11)$$

According to (11), the median is the maximum likelihood estimate (MLE) of location for the Laplacian distribution:

$$f(x) = \frac{1}{2} e^{-|x|}. \quad (12)$$

Traditionally, the median filter performance is compared to the performance of the *moving average filter*:

$$y_i = \frac{1}{n} \sum_{j=i-\nu}^{i+\nu} x_j, \quad (13)$$

which is essentially a “moving” arithmetic mean. The arithmetic mean is the MLE estimate of location for the Gaussian distribution and it minimizes the L_2 norm. The asymptotic relative efficiency of the median versus the arithmetic mean $ARE(\text{med}(x_i), \bar{x})$ (4) is shown in Table 1. It can easily be seen that the median is more efficient for long-tailed distributions, whereas the arithmetic mean is more efficient for the Gaussian distribution and for short-tailed distributions. This means that the moving average filter performs better than the median filter in additive Gaussian noise removal [5], [25]. The $ARE(\text{med}(x_i), \bar{x})$ for various other distributions can be found in [5] and [9].

The median is a B-robust operator, because its influence function is bounded [7]:

$$IF(x; \text{med}, F) = \frac{1}{2f(F^{-1}(1/2))} \text{sign}(x - F^{-1}(\frac{1}{2})). \quad (14)$$

Therefore, a single outlier (e.g. impulse) can have no effect on its performance, even if its magnitude is very large or very small. The influence function of the arithmetic mean for the Gaussian distribution is given by [7]

$$IF(x; \bar{x}, F) = x \quad (15)$$

and it is unbounded. Therefore, the moving average filter is very susceptible to impulses. The breakdown point of the median is $\epsilon^* = 1/2$. The median becomes unreliable only if more than 50% of the data are outliers. The arithmetic mean has $\epsilon^* = 0$. Even a single outlier can destroy its performance.

The statistical and robustness properties of the median make it very suitable for impulsive noise filtering. The performance of the median in impulsive noise filtering can be measured by the probability of correct signal reconstruction. If the impulses have constant value and probability

Table 2 Rate of Failure of the Median Filter in Removing Impulse Noise

Impulse Rate p	Window Size n			
	3	5	9	25
0.1	0.028	0.0086	0.00089	0.000002
0.2	0.104	0.058	0.0196	0.00037
0.3	0.216	0.163	0.099	0.017
0.4	0.352	0.317	0.267	0.154
0.5	0.5	0.5	0.5	0.5

of existence p in a constant image neighborhood, the probability of correct reconstruction, $P(n, p)$, is given by [25]

$$P(n, p) = \sum_{k=0}^{(n-1)/2} \binom{n}{k} p^k (1-p)^{n-k}. \quad (16)$$

The probabilities of erroneous reconstruction, $1 - P(n, p)$, are given in Table 2. It can be seen that a 3×3 median filter ($n = 9$) can reject impulses having 30% probability of occurrence with probability of success over 90%. Figure 2 illustrates that the median removes impulses efficiently, whereas the moving average fails to do so.

The median filter has low-pass characteristics if its input is white additive noise, as can be found by examining the autocorrelation function and the power spectrum of its output [5], [28]. The autocorrelation function can be calculated numerically by the joint pdf of the filter output [28]–[30]. In the case of nonwhite input noise, the evaluation of median filter performance is much more difficult. Its performance is certainly inferior to that of the moving average filter [25].

As stated in the introduction, edge information is very important for human perception. Therefore, its preservation and, possibly, its enhancement constitute a very important subjective feature of the performance of an image filter. Edges, by definition, contain high frequencies. Therefore, low-pass linear filters, e.g. the moving average filter, smooth them and produce images which are unpleasant to the eye. In contrast, the median filter tends to preserve the edge sharpness, owing to its robustness properties. An idealized noisy vertical edge model of height h is shown in Fig. 3(a). It is described by the following equation:

$$x_{ij} = \begin{cases} z_{ij} & j \leq 0 \\ h + z_{ij} & j \geq 1, \end{cases} \quad (17)$$

where h denotes the edge height and z_{ij} is white noise. It is clearly seen, by comparing parts (b) and (c) of Fig. 3, that the median filter preserves edges much better than the moving average filter. Figure 4 shows an image corrupted by additive Gaussian noise. The 7×7 median filter produces a much more pleasant output than that of a 7×7 moving average filter, because it does not blur edges. However, the median filter output contains more

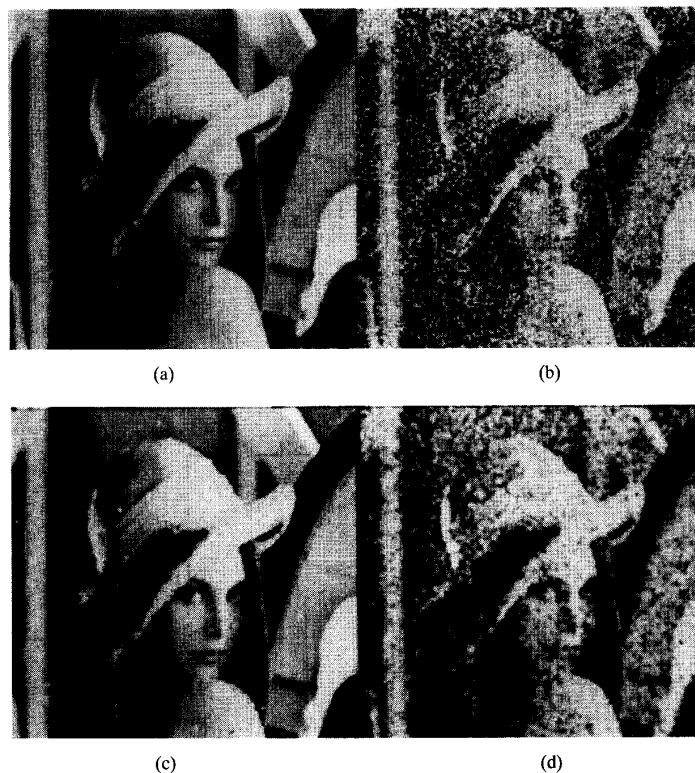


Fig. 2. (a) Original image; (b) image corrupted by salt-pepper noise; (c) 3×3 median filter output; (d) 3×3 moving average filter output.

noise in the homogeneous image regions than the moving average filter output. A quantitative analysis of the edge preservation properties of the median filter can be found in [32]–[34].

The median filter not only smooths noise in homogeneous image regions; it also tends to produce regions of constant or nearly constant intensity [35]. The shape of these regions depends on the geometry of the filter window. Usually, they are either linear patches (streaks) or blotches. These effects are undesirable because they are perceived as lines or contours, which do not exist in the original image. A theoretical analysis of the streaking effect can be found in [35]. Rabiner *et al.* [36] proposed postprocessing of the median filter output by short linear smoothers. However, if streaking is still a serious problem, the only solution is to use other order statistics filters, to be described later on in this section.

B. Deterministic Analysis

A very important tool in the analysis of linear systems is their steady-state behavior with sinusoidal inputs. It is known that the frequency of a sinusoid is not changed when it passes through a linear system. Only its phase and amplitude are changed. Of course, this is not valid for median filters, because they are highly nonlinear systems. Therefore, many researchers have focused on the problem of finding signals (called roots or fixed points), which are

invariant under median filtering [37]–[39]. Median roots have found applications in speech and image coding [38], [43], [49]. Several problems are related to the median roots. The first one is the determination of the shape of a signal which is a root of a one- or two-dimensional median filter. The second problem concerns the construction and counting of the number of median roots. The third problem is the rate of convergence of a nonroot signal to a root after successive passes through the median. These three problems form the subject of the deterministic analysis of median filters. Certain concepts which are used in the deterministic analysis of a median filter of length $n = 2\nu + 1$ have to be defined.

- A sequence x_i is *monotonic* if $x_i \leq x_j$ or $x_i \geq x_j$ for every $i < j$.
- A sequence x_i is *locally monotonic* of length m if the subsequence (x_i, \dots, x_{i+m-1}) is monotonic for every i .
- A *constant neighborhood* consists of at least $\nu + 1$ consecutive identically valued points.
- An *edge* is a monotonic region between two constant neighborhoods having different values.

If a signal x_i is a root of a median filter of length $n = 2\nu + 1$ and has a monotonic segment of length ν , x_i is locally monotonic of length $\nu + 2$ [37]. This practically means that, with the exception of certain periodic signals, a root consists of constant neighborhoods and edges. It has also

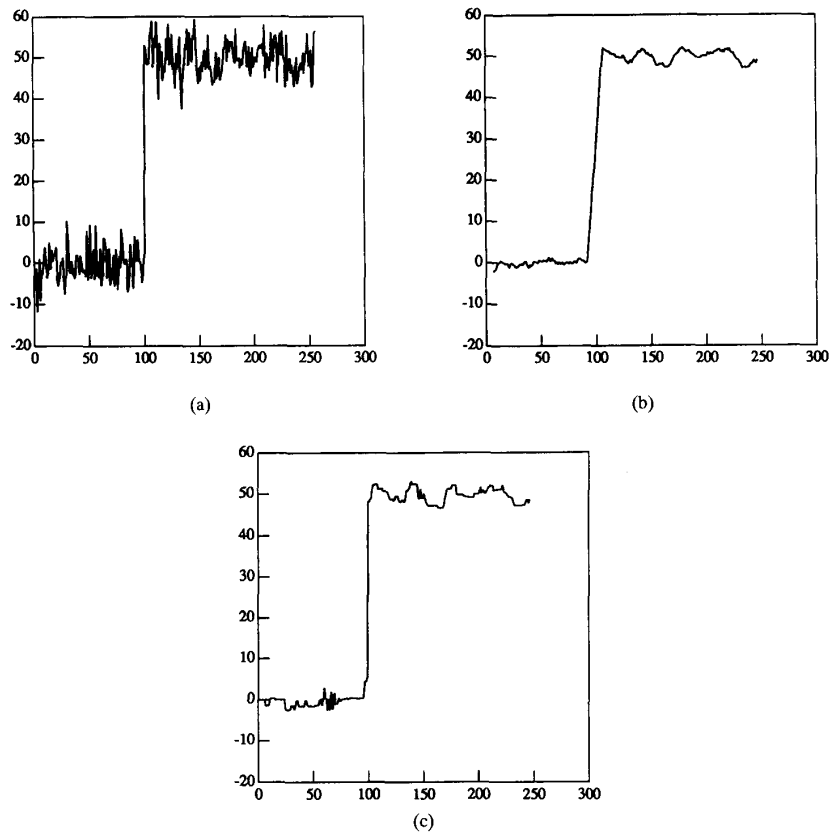


Fig. 3. (a) Noisy image edge. (b) Its degradation by a moving average filter. (c) Median filter output.

been shown that, for finite-length signals, a signal is a root if and only if it consists of constant neighborhoods and edges [39]. The patterns of the roots of two-dimensional medians are much more complicated. The interested reader can find more details about their shape in [37], [40]. Locally monotonic signals and their linear regression have been proposed for digital signal smoothing [41], [42].

If k -valued signals of finite length L are considered, the number of median roots is finite and depends on k , L . Therefore, they can be counted and constructed in a deterministic way [43], [46]. In the following, binary signals ($k = 2$) will be considered, because it will be shown later on that the median of a k -valued signal can be decomposed to the median of binary signals. Let us assume that the median size is $n = 3$ ($\nu = 1$). The binary root signal of such a median consists of constant neighborhoods of minimal length 2, i.e., (00), (11) and of transitions (edges) of the form (01), (10). The signal sequence x_1, x_2, \dots is padded by the sample x_0 at its starting point. If the first sample, x_1 , is assumed to be zero, the appended sample x_0 has to be 0 too. They already form a constant neighborhood (00). The second sample, x_2 , can be either 1 (forming an edge) or 0 (continuing the constant neighborhood). Therefore, $(x_0 x_1 x_2)$ can have two allowable states (000) or (001). This process can be

repeated until sequences of length L are constructed. It can be described by the tree structure, shown in Fig. 5. In this figure each path represents a signal. The nonroot signals are denoted by X. Therefore, by counting the allowable paths, the number of roots can be found. The number of roots of a three-point median, $n = 3$, which have length L is given by [43]

$$f(L) = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^L - \left(\frac{1 - \sqrt{5}}{2} \right)^L \right]. \quad (18)$$

As expected, the number of roots increases very fast with their length L . A similar formula can be found for k -valued roots of medians having arbitrary window size n . Their number increases tremendously with k and n [44], [45].

Each finite length nonroot signal must converge to a root signal after consecutive passes with a median filter [39]. It has been proved [47] that a nonroot signal will converge to a root in at most $3 \lceil \frac{L-2}{2(\nu+2)} \rceil$ passes of a median of length $n = 2\nu + 1$. Usually, fewer passes are required [45]. Convergence theorems are presented in [48]. An example of a three-level signal that converges to a root is shown in Fig. 6. Two median passes are required in this case.

Several modifications and extensions of the median filters have been proposed in the literature. Some of them (e.g. stack filters) have rigorous mathematical foundations and

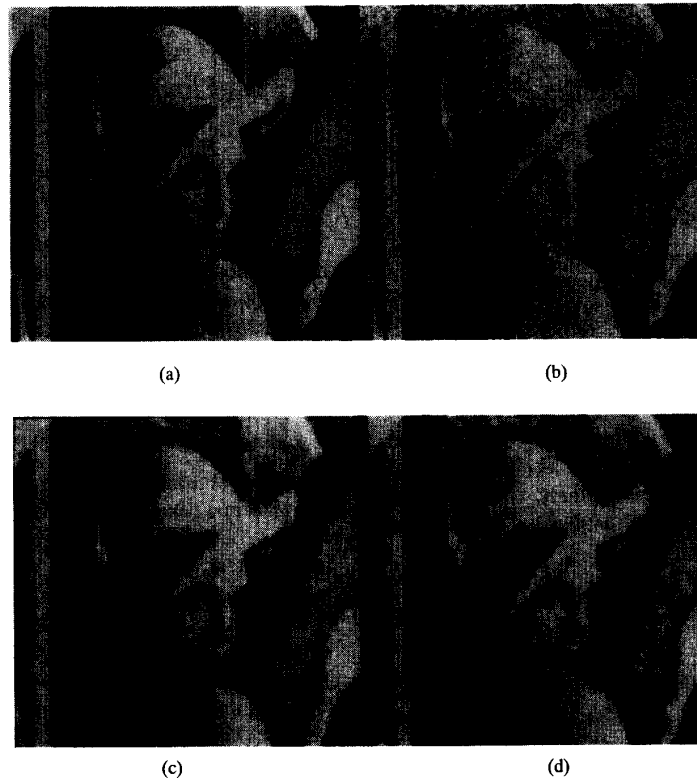


Fig. 4. (a) Original image. (b) Image corrupted by additive white Gaussian noise having standard deviation 100. (c) Output of the 5×5 moving average filter. (d) Output of a 5×5 median filter.

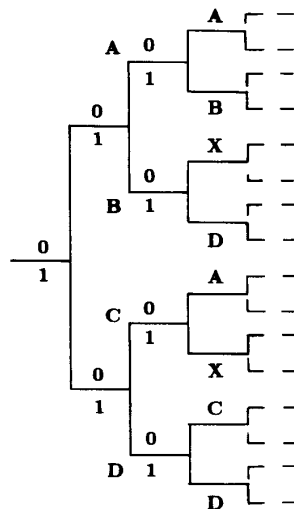


Fig. 5. Tree structure for the binary roots of a median filter of length 3.

provide a unifying framework for the design of filters based on order statistics. Other (e.g. separable median filters) aim at the reduction of the computational complexity for median filter computation. Finally, certain filters (e.g. multistage median filters, median hybrid filters) have been designed

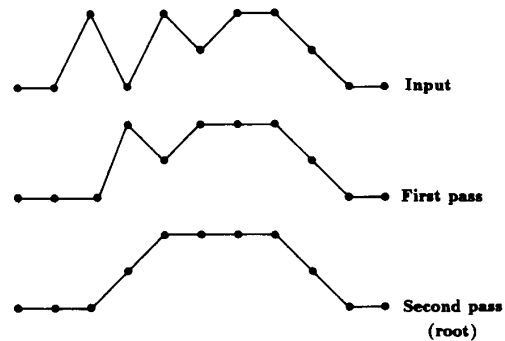


Fig. 6. A nonroot signal and its convergence to a root after two consecutive median passes.

to improve the image detail preservation properties of the median filter. In the following we shall give a brief description of these filters.

1) *Threshold Decomposition and Stack Filters:* Median filtering of binary signals is relatively easy and fairly well understood. Its computation can be reduced to counting the 1's inside the filter window. If their number is greater than or equal to $\nu + 1$, the output of the median is 1; otherwise it is 0. Furthermore, the properties of binary signals in median filtering are fairly well understood. Therefore, it is attractive from both a practical and a theoretical point

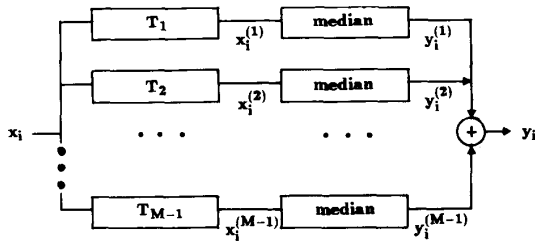


Fig. 7. Threshold decomposition of a median filter.

of view to reduce the median filtering of k -valued signals to median filtering of binary signals. This can be obtained by the *threshold decomposition* of the median filter [50]. Let x_i be an M -valued signal: $0 \leq x_i \leq M - 1$, and consider the $M - 1$ thresholds: $1 \leq j \leq M - 1$. The signal x_i can be decomposed to $M - 1$ binary valued signals $x_i^{(j)}$, $1 \leq j \leq M - 1$ by using a threshold decomposition function $T_j(x_i)$ as follows:

$$x_i^{(j)} = T_j(x_i) = \begin{cases} 1 & \text{if } x_i \geq j \\ 0 & \text{if } x_i < j. \end{cases} \quad (19)$$

These $M - 1$ binary valued signals can be filtered independently:

$$y_i^{(j)} = \text{med}(x_{i-\nu}^{(j)}, \dots, x_i^{(j)}, \dots, x_{i+\nu}^{(j)}), \quad 1 \leq j \leq M - 1. \quad (20)$$

The M th valued median output y_i can be reconstructed by summing the binary output signals:

$$y_i = \sum_{j=1}^{M-1} y_i^{(j)}. \quad (21)$$

A schematic diagram of the threshold decomposition of median filters is shown in Fig. 7. Threshold decomposition has both theoretical and practical significance for median filtering [50], [51]. It has been used to find the statistical distribution of the output of the median filter and of the roots of the median filter [51], [52]. It has also been used to develop architectures for fast median filtering [53].

Let us suppose that we denote the vector of the binary signals $x_i^{(j)}$ and the vector of multivalued signals x_i by $\mathbf{x}_i^{(j)} = [x_{i-\nu}^{(j)}, \dots, x_{i+\nu}^{(j)}]^T$ and $\mathbf{x}_i = [x_{i-\nu}, \dots, x_{i+\nu}]^T$ respectively. We denote by $f(\mathbf{x})$ a binary operator that operates on a binary vector of length n . We decompose the vector \mathbf{x}_i by threshold decomposition and we apply the binary operator f at each level $j = 1, \dots, M - 1$ independently. The outputs of these binary filters possess what is referred to as the stacking property if the binary output signals $y_i^{(j)}$, $1 \leq j \leq k - 1$ are piled and the pile of y_i at time i consists of a column of 1's having a column of 0's on top. The filters that support the stacking property are called stack filters. Their structure is similar to the one shown in Fig. 7, where the binary medians are replaced by the binary operators f . Stack filters possess both the threshold decomposition structure and the stacking

property. The output of a stack filter is given by

$$y_i = S_f(\mathbf{x}_i) = \sum_{j=1}^{M-1} y_i^{(j)} = S_f\left(\sum_{j=1}^{M-1} T_j(\mathbf{x}_i)\right) = \sum_{j=1}^{M-1} S_f(T_j(\mathbf{x}_i)) = \sum_{j=1}^{M-1} f(\mathbf{x}_i). \quad (22)$$

Stack filters constitute a broad class of nonlinear filters having median filters as a special case. For example, the binary median filter operating on a binary signal x_i is essentially a Boolean function f of the binary variables $x_{i-\nu}, \dots, x_i, \dots, x_{i+\nu}$. Its form is given by the following relation for the filter length $n = 3$ [53]:

$$y_i = x_{i-1}x_i + x_i x_{i+1} + x_{i+1}x_{i-1}, \quad (23)$$

where $., +$ denote Boolean operators. Such a binary function operates in (20) on each binary signal $x_i^{(j)}$, $1 \leq j \leq M - 1$ obtained by the threshold decomposition. This remark greatly facilitates the VLSI implementation of the median filter based on threshold decomposition [53].

The class of stack filters is very large. Their number, when the filter window is n , grows faster than $2^{2^{n/2}}$ [58]. The problem of finding them reduces to the problem of finding *stackable* Boolean functions, because it is the Boolean functions that determine the properties of the decomposed filter. It has been established that a Boolean function f is stackable if and only if it contains no complements of the input variables of the form \bar{x}_i . Thus, the stackable functions are also called positive Boolean functions. The binary median (23) is stackable because it does not contain complements of any of the inputs x_{i-1}, x_i, x_{i+1} . The number of the positive Boolean functions is finite for finite filter length n . There are, respectively, 20 and 7581 such functions for $n = 3$ and 5. Examples of stackable functions for $n = 3$ are the following:

$$y_i = x_i + x_{i-1}x_{i+1} \quad (24)$$

$$y_i = x_{i-1}x_i + x_i x_{i+1}. \quad (25)$$

The stack filters based on (24) and (25) are called *asymmetric medians* because they eliminate negative and positive impulses, respectively. Any composition of the stack filters can be expressed by a single stack filter, due to a similar property of the positive Boolean functions. Generalized stack filters have also been proposed [56], [57]. Such filters allow more than one binary signals to be fed to the binary filters f . Furthermore, different binary filters f_i are allowed at the various thresholding levels. Recursive stack filters have also been proposed and their properties have been analyzed [60].

Great advances have been made recently in the design of stack filters [54]–[59]. An estimation approach and a structural approach have been proposed [57]. The estimation approach employs the *minimum absolute error (MAE)* criterion because of its robustness [54], [55]. Let us suppose that a signal s_i is corrupted by noise and that the observed

signal x_i is produced. This signal is fed to a stack filter S_f whose output is y_i . The MAE is given by

$$E = E[|s_i - y_i|] = E[|s_i - S_f(\mathbf{x}_i)|] = E\left[\left|\sum_{j=1}^{M-1} (s_i^{(j)} - S_f(T_j(\mathbf{x}_i)))\right|\right] = E\left[\left|\sum_{j=1}^{M-1} (s_i^{(j)} - f(\mathbf{x}_i^{(j)}))\right|\right]. \quad (26)$$

It can be proved [55] that the MAE takes the form

$$E = \sum_{j=1}^{M-1} E[|s_i^{(j)} - f(\mathbf{x}_i^{(j)})|] \quad (27)$$

and it is decomposed to the sum of errors at each threshold level. Based on this reduction, a binary function f can be chosen in such a way so that it minimizes a cost function. Let \mathbf{z}_j be a binary vector of length n . The output of the binary function $f(\mathbf{z}_j)$ is called the decision variable $P_f(1|\mathbf{z}_j)$. It can take binary values $P_f(1|\mathbf{z}_j) = 0, 1$. The Boolean function f can be represented as a vector of length 2^n . Its j th entry is $P_f(1|\mathbf{z}_j)$. Let us suppose that we denote by C_j the cost incurred when the function $f(\mathbf{z}_j)$ produces output 1. It can be shown that the MAE (27) can take the form

$$E = \sum_{j=1}^{2^n} C_j P_f(1|\mathbf{z}_j). \quad (28)$$

This cost function must be minimized with respect to $P_f(1|\mathbf{z}_j)$ under the stacking constraint:

$$P_f(1|\mathbf{z}_j) \leq P_f(1|\mathbf{z}_k) \quad \text{if } \mathbf{z}_j \leq \mathbf{z}_k \quad (29)$$

and the zero-one constraint:

$$P_f(1|\mathbf{z}_j) = 0, 1. \quad (30)$$

The notation $\mathbf{z}_j \leq \mathbf{z}_k$ means that each element of \mathbf{z}_j is less than or equal to the corresponding element of \mathbf{z}_k . This is a zero-one integer linear programming problem. It can be transformed to a linear programming problem [55], [59]:

$$\text{minimize } \sum_{j=1}^{2^n} C_j P_f(1|\mathbf{z}_j) \quad (31)$$

under the constraints

$$P_f(1|\mathbf{z}_j) \leq P_f(1|\mathbf{z}_k) \quad \text{if } \mathbf{z}_j \leq \mathbf{z}_k \\ 0 \leq P_f(1|\mathbf{z}_j) \leq 1. \quad (32)$$

The quantity $P_f(1|\mathbf{z}_j)$ can be interpreted as the probability that the filter f will produce output 1 when its input is \mathbf{z}_j . The filter can be obtained by optimizing either (28)–(30) or (31) and (32). The resulting solutions are not necessarily unique. The computational complexity of the optimization is of the order $O(n2^n)$, whereas the direct search of the optimal solution has complexity of the order $O(2^{2^n/2})$. The above-mentioned design technique has also been used for the design of generalized stack filters [56]. Structural constraints and goals [57] have also been incorporated in the optimization problem [58]. Finally, adaptive versions of stack filters have been defined by using the MAE [59].

2) *Separable Median Filter*: A separable two-dimensional median filter of size n results from two successive applications of a one-dimensional median filter of length n along rows and then along columns of an image (or vice versa):

$$y_{ij} = \text{med}(z_{i,j-\nu}, \dots, z_{ij}, \dots, z_{i,j+\nu}) \quad (33)$$

$$z_{ij} = \text{med}(x_{i-\nu,j}, \dots, x_{ij}, \dots, x_{i+\nu,j}). \quad (34)$$

The separable median filter of length n has greater output variance than a nonseparable one having $n \times n$ extent [61]. Its main advantage is its low computational complexity in comparison with that of the nonseparable median filter, because it sorts n numbers two times, whereas the nonseparable $n \times n$ median sorts n^2 numbers. The deterministic analysis of the separable median filters can be found in [62] and [63].

3) *Recursive Median Filter* One intuitive modification of the median filters is to use the already computed output samples $y_{i-\nu}, \dots, y_{i-1}$ in the calculation of y_i :

$$y_i = \text{med}(y_{i-\nu}, \dots, y_{i-1}, x_i, \dots, x_{i+\nu}). \quad (35)$$

This filter is a recursive median filter. Its output tends to be much more correlated than that of the standard median filter. This property is easily explained by the fact that the output samples $y_{i-\nu}, \dots, y_{i-1}$ contribute directly to the computation of y_i . Recursive median filters have higher immunity to impulsive noise than nonrecursive median filters [64]. Furthermore, they have some nice deterministic properties [62]. A signal is invariant to recursive median filtering if and only if it is invariant to standard median filtering. Furthermore, any nonroot signal is reduced to a root signal after just one pass by a recursive median filter. However, this root is not necessarily the same as the root which is obtained by standard median filtering. Fast recursive algorithms for the calculation of two-dimensional median roots can be found in [65]. A variation of the recursive median filter is the separable recursive median filter:

$$y_{ij} = \text{med}(y_{i,j-\nu}, \dots, y_{i,j-1}, \dots, z_{ij}, \dots, z_{i,j+\nu}) \quad (36)$$

$$z_{ij} = \text{med}(z_{i-\nu,j}, \dots, z_{i-1,j}, \dots, x_{ij}, \dots, x_{i+\nu,j}). \quad (37)$$

Recursive separable median filters perform better than the nonrecursive ones in impulsive noise removal [66]. A deterministic analysis of the separable recursive filters can be found in [67].

4) *Weighted Median Filters*: The *weighted median* is the estimator that minimizes the weighted L_1 norm of the form

$$\sum_{i=1}^n w_i |x_i - T_n| \rightarrow \min. \quad (38)$$

Therefore it has close relation to the standard median according to (11). It can be proved that the minimization

of (43) leads to the following explicit form of the weighted median [69]:

$$T_n = \text{med}(w_1 \square x_1, \dots, w_n \square x_n), \quad (39)$$

where $w \square x$ denotes duplication of x w times:

$$w \square x = x, \dots, x \quad (w \text{ times}). \quad (40)$$

The sum of the weights $w_i, i = 1, \dots, n$, must be an odd number. The weighted median filter was first introduced by Justusson [25]:

$$y_i = \text{med}(w_{-\nu} \square x_{i-\nu}, \dots, w_{\nu} \square x_{i+\nu}). \quad (41)$$

It is closely related to the FIR filter having the form

$$y_i = \frac{\sum_{j=-\nu}^{\nu} w_j x_{i+j}}{\sum_{j=-\nu}^{\nu} w_j}. \quad (42)$$

The analysis of the performance of the weighted median filter can be found in [69]–[71]. It is shown that it can outperform the median filter [71]. The connection between stack filters and weighted median filters is described in [70]. This connection is used to derive the statistical and deterministic properties of weighted median filters. Weighted median filters have been applied to image filtering and to speckle noise removal in ultrasound images [72], [73]. Finally, weighted median filters have been used in spatiotemporal image sequence filtering [77].

5) *Max-Median Filters and Multistage Median Filters:* Several rather unrealistic assumptions are made in the theoretical analysis of the median filters (especially in the deterministic case), e.g. that the image consists of constant neighborhoods and edges. In reality images have fine details, such as lines and sharp corners, which are very valuable for human vision. These details are usually destroyed by medians having relatively large windows (larger than 5×5). It is the rank ordering process that destroys the structural and spatial neighborhood information. Thus, several efforts have been made to take into account spatial information in median filter design. Such a modification of the median leads to the max-median filters [74]:

$$y_{ij} = \max(z_1, z_2, z_3, z_4) \quad (43)$$

$$z_1 = \text{med}(x_{i,j-\nu}, \dots, x_{ij}, \dots, x_{i,j+\nu}) \quad (44)$$

$$z_2 = \text{med}(x_{i-\nu,j}, \dots, x_{ij}, \dots, x_{i+\nu,j}) \quad (45)$$

$$z_3 = \text{med}(x_{i+\nu,j-\nu}, \dots, x_{ij}, \dots, x_{i-\nu,j+\nu}) \quad (46)$$

$$z_4 = \text{med}(x_{i-\nu,j-\nu}, \dots, x_{ij}, \dots, x_{i+\nu,j+\nu}). \quad (47)$$

The subsections of the max/median filter span the vertical, the horizontal, and the two diagonal lines that pass through the point (i, j) . Thus, spatial information is taken

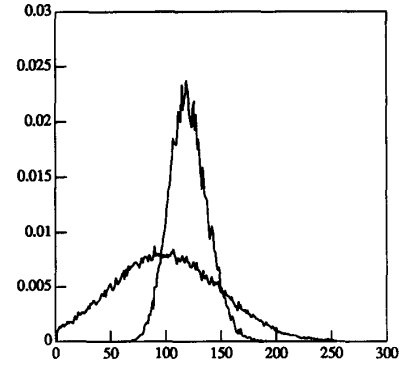


Fig. 8. Output distributions of the max/median filter (high curve), when the input distribution is Gaussian (low curve).

into account by the max/median filter. The max operator produces biasedness toward high intensity levels. Therefore, the max/median filter is a biased estimator of the mean, as can be seen in Fig. 8 [74], [75]. Its performance can be improved considerably if the median operator is used to replace the max operator in (43). The resulting filter belongs to what are referred to as multistage median filters. Such filters are described by the following relation:

$$y_{ij} = \text{med}(\text{med}(z_1, z_2, x_{ij}), \text{med}(z_3, z_4, x_{ij}), x_{ij}). \quad (48)$$

Multistage median filters can preserve details in horizontal, diagonal and vertical directions, because they use subfilters that have regions of support along these directions. Their theoretical analysis can be found in [76]. The use of multistage median filters in moving image processing has been described in [77], [78]. It has been proven that spatiotemporal multistage filtering outperforms both spatial and temporal filtering in terms of noise suppression. Other modifications of the max/median filter can be found in [79].

6) *Median Hybrid Filters:* Hybrid filters are a combination of linear filters and median filters. The main purpose of using linear subfilters is to take into account spatial information about the image. The median hybrid filters have the following definition [80]:

$$y_i = \text{med}(\Phi_1(x_i), \dots, \Phi_M(x_i)), \quad (49)$$

where the filters $\Phi_j(x_i), j = 1, \dots, M$ are linear FIR or IIR filters. The structure of such a median hybrid filter, using FIR subfilters, is shown in Fig. 9. A special case of an FIR hybrid median filter is the following:

$$y_i = \text{med}\left(\frac{1}{\nu} \sum_{j=1}^{\nu} x_{i-j}, x_i, \frac{1}{\nu} \sum_{j=1}^{\nu} x_{i+j}\right). \quad (50)$$

The performance of this median hybrid filter in a noisy step edge is shown in Fig. 10. The sections Φ_1, Φ_3 lag and lead the edge respectively, whereas Φ_2 reacts at the edge. Thus spatial information is incorporated in the performance of the median filter and the edge is preserved. The performance of median hybrid filters is analyzed in [80], [81], and [83].

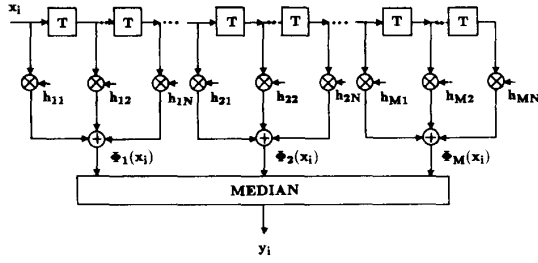


Fig. 9. Structure of an FIR median filter.

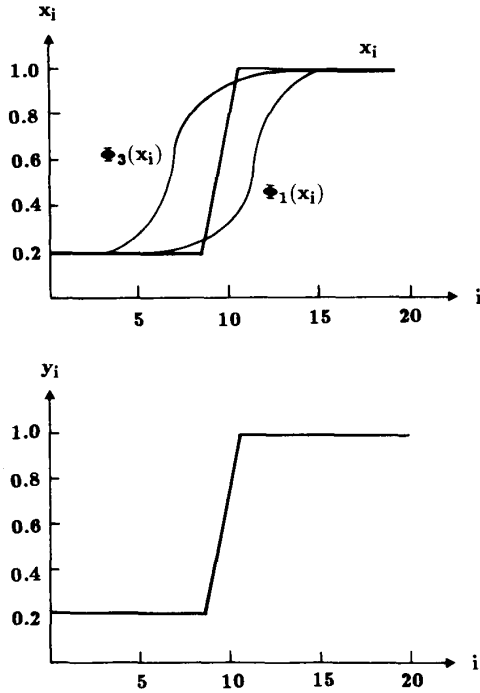


Fig. 10. The performance of a median hybrid filter in the presence of an edge.

Modifications of the FIR median hybrid filters with good transient response are presented in [86]. FIR median hybrid filters with predictive FIR substructures have also been designed [82]. The corresponding recursive median hybrid filter is a modification of the filter (50):

$$y_i = \text{med}\left(\frac{1}{\nu} \sum_{j=1}^{\nu} y_{i-j}, x_i, \frac{1}{\nu} \sum_{j=1}^{\nu} x_{i+j}\right). \quad (51)$$

The hybrid median filters can be easily extended to two dimensions [84]. Such a filter is described by the following equation:

$$y_{ij} = \text{med}(z_{Nij}, z_{Sij}, z_{Eij}, z_{Wij}, x_{ij}), \quad (52)$$

where z_{kij} , $k = N(\text{orth}), S(\text{outh}), E(\text{ast}), W(\text{est})$ are FIR filters whose region of support is shown in Fig. 11. The

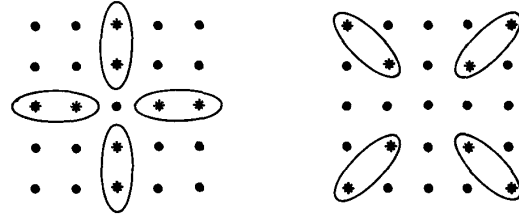


Fig. 11. Regions of support for two-dimensional FIR median hybrid filters.

region of support of a version of the same filter rotated by 45° is shown in the same figure. Two-dimensional median hybrid filters preserve fine image details fairly well [84]. However the output variance of a median hybrid filter (50) is approximately equal to that of a moving average filter of half the extent [81]. Therefore, it is less immune to additive white noise than the moving average filter. A VLSI implementation of hybrid median filter is described in [85].

III. FILTERS BASED ON ORDER STATISTICS

The class of filters based on order statistics is very rich. Besides the median filter and its modifications, it includes a large number of nonlinear filters. Some of them (e.g. the ranked-order filters and the max/min filters) are straightforward applications of order statistics in filtering. Other filters (e.g. α -trimmed mean filters, L-filters) come from the application of robust estimation techniques in digital signal/image filtering. They are closely related to a large class of robust estimators called *L estimators* [7]–[9]. This section gives a review of nonlinear filters based on order statistics. It also contains a brief description of some other nonlinear filters that stem from classes of robust estimators, namely from *L estimators* and *R estimators*. It will be seen that these filters are related to filters based on order statistics.

A. Ranked-Order Filters

An r th ranked-order filter of the signal x_i is the r th order statistic:

$$y_i = r\text{th order statistic of } \{x_{i-\nu}, \dots, x_i, \dots, x_{i+\nu}\} \quad (53)$$

of the signal data within the filter window. The probability distribution of the output of a rank order filter has already been described in Section II. It can be found, by examining the output pdf (10) and the output mean, that the r th rank order filter introduces a strong bias to the estimation of the mean, when the rank is small or large. In this case, the filter tends to perform as a maximum or minimum filter respectively. The bias is even stronger when the input data have a long-tailed distribution. For long-tailed distributions, the output variance attains a minimum when the r th rank is close to the median. This fact is explained by the good performance of the median in long-tailed distributions, as has already been stated in Section II. For

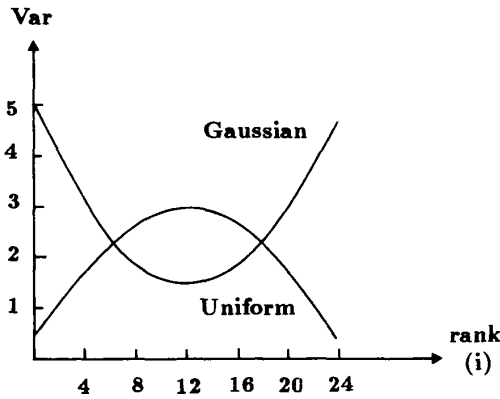


Fig. 12. Variance of the output of the r th rank order filter.

short-tailed distributions, r th ranks that are close to min or to max give lower output variance. This behavior is illustrated in Fig. 12, where the variance of the r th-order statistic is given as a function of r for uniform and Gaussian distributions of identically distributed input data [87]. It is clear from the previous analysis that the rank r must be chosen carefully according to the input distribution and the desired filter output. Ranked-order filters have been applied as postdetection processor in radar detection systems [88]. Such filters have been shown to be consistent and biased estimators of the received signal distributions. Threshold decomposition, which has been described in the previous section, can also be applied to ranked-order filters [89]. It can be used both for their theoretical analysis and for their implementation. Such VLSI implementations have already been reported [90]–[92]. Ranked-order filters using pixel neighborhoods different from that of (53) can be found in [93]. The relation between ranked-order filters and the mean absolute error criterion is described in [54].

B. Max/Min Filters

The maximum $x_{(n)}$ and the minimum $x_{(1)}$ are the two extremes of the rank order filters. They are closely related to two morphological operations, called, respectively, dilation and erosion, of a function by a set [5], [94]–[104]. This review paper cannot cover the entire discipline of mathematical morphology. Therefore, this section will cover only certain aspects of the relation between mathematical morphology and order statistics filters. This relation has been analyzed in [105] and [106]. It has been proved that any ranked-order filter can be represented as a max or min superposition of erosions or dilations respectively [105]. The maximum filter can effectively remove negative impulses (black spots) in an image, whereas the minimum filter can remove positive impulses (white spots). Both fail in the removal of mixed impulsive noise, because minimum and maximum filters tend to enhance the negative and positive spikes respectively. However, cascades of max and min filters can effectively remove such a mixed impulsive noise. A maximum filter followed by a minimum filter is the mor-

phological filter called the closing filter [94]. A minimum filter followed by a maximum filter is the morphological filter called the opening filter [94]. A closing filter followed by an opening filter forms a *close-opening (CO)* filter. An opening filter followed by a closing filter forms an *open-closing (OC)* filter [107]. Both close-opening and open-closing filters perform relatively well in the case of impulsive noise. However, their performance is generally inferior to that of the median filter [5]. Furthermore, they are biased estimators of location in the case of additive white noise [107]. Multiscale morphological operators have also been proposed recently for nonlinear noise smoothing [108]. Both the maximum and the minimum filters have good edge preservation properties. Their disadvantage is that they tend to enhance the bright and the dark regions of the image respectively [5], as can be seen in Fig. 13. The maximum and minimum filters are very popular, especially in the context of mathematical morphology, despite their disadvantages. The main reason is their computational simplicity and the existence of fast processors for their calculation [103]. The maximum/minimum filters are related to the L_p -mean filters [109] given by

$$y_i = \left(\sum_{j=1}^n a_j x_{i-\nu+j-1}^p \right)^{1/p}, \quad p \in \mathbf{Z}. \quad (54)$$

The L_p -mean filter tends to the maximum filter when p tends to infinity. The L_{-p} -mean filter tends to the minimum filter when p tends to infinity. Therefore, if the effects arising from the highly nonlinear nature of max and min operators are not desirable (e.g. the enhancement of dark or bright regions), L_p -mean or L_{-p} -mean filters of moderate p can be used instead. L_p -mean filters have relatively good noise smoothing and edge preservation properties [109].

C. α -Trimmed Mean Filters

It has already been stated that the moving average filter suppresses additive white Gaussian noise better than the median filter, whereas the second is better at preserving edges and rejecting impulses. Therefore, a good compromise between the two is highly desirable. Such a filter is the α -trimmed mean filter [110]:

$$y_i = \frac{1}{n(1-2\alpha)} \sum_{j=\alpha n+1}^{n-\alpha n} x_{(j)}, \quad (55)$$

where $x_{(j)}$, $j = 1, \dots, n$ are the order statistics of $x_{i-\nu}, \dots, x_i, \dots, x_{i+\nu}$. The α -trimmed mean filter rejects the smaller and the larger observation data. Data rejection depends on the coefficient α , $0 \leq \alpha < 0.5$. If α is zero, no data are rejected and the filter performs as a moving average filter. If α is close to 0.5, all data but the median are rejected. Therefore, the α -trimmed mean filter can be used as a compromise between the median filter and the moving average filter. The rejected data are usually outliers. This fact explains the robustness of the α -trimmed mean filter [7]. Its robustness is controlled by the parameter α . Its

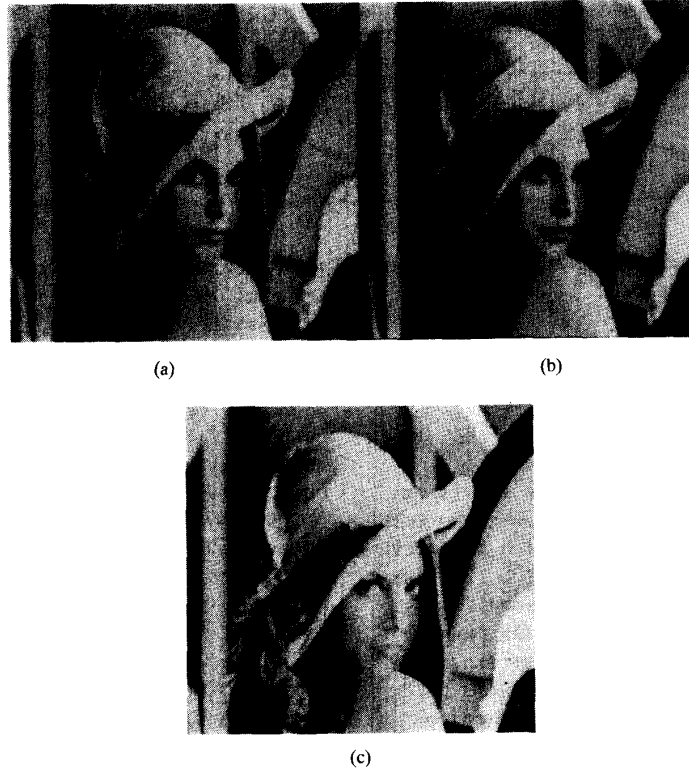


Fig. 13. (a) Original image. (b) Min filter output. (c) Max filter output.

breakdown point, ϵ^* , is equal to α . Therefore it can reject up to $100\alpha\%$ of outliers. It has also been proved that the α -trimmed mean has better performance than the moving average filters for long-tailed distributions [9]. However, for short-tailed distributions, its performance is poor. It is known that the *midpoint MP*:

$$MP = \frac{1}{2}(x_{(1)} + x_{(n)}) \quad (56)$$

is a very good estimator of location for short-tailed distributions (e.g. uniform distribution) [9], [111]. Therefore, the so-called complementary α -trimmed mean, which is close to the midpoint, has been proposed for short-tailed input noise distributions [110]:

$$y_i = \frac{1}{2\alpha} \left[\sum_{j=1}^{\alpha n} x_{(j)} + \sum_{j=n-\alpha n+1}^n x_{(j)} \right]. \quad (57)$$

A different approach to trimmed filters is to exclude the samples $x_{i+r,j+s}$ in the filter window, which differ considerably from the local median $med(x_{ij})$. This is the *modified trimmed mean (MTM)* filter [112]:

$$y_{ij} = \frac{\sum \sum_A a_{rs} x_{i+r,j+s}}{\sum \sum_A a_{rs}}. \quad (58)$$

The summations cover the entire filter window A . The filter

coefficients are chosen as follows:

$$a_{rs} = \begin{cases} 1 & \text{if } |x_{i+r,j+s} - med(x_{ij})| \leq q \\ 0 & \text{otherwise.} \end{cases} \quad (59)$$

The amount of trimming depends on the parameter q . Data deviating strongly from the local median are trimmed out. Since such data are usually outliers, the modified trimmed mean filter has good robustness properties. A variation of (58) and (59) employs two filter windows of different sizes and is called the double window modified trimmed mean (DW MTM) filter [112]. It is known to have good robustness and edge preservation properties. Another modification is the *modified nearest neighbor (MNN)* filter [113], whose coefficients are given by

$$a_{rs} = \begin{cases} 1 & \text{if } |x_{i+r,j+s} - x_{ij}| \leq q \\ 0 & \text{otherwise.} \end{cases} \quad (60)$$

This filter trims out pixels deviating strongly from the central pixel. Therefore, MNN filters have good edge preservation properties.

D. L Filters An important generalization of the median is the L filter (also called the order statistic filter):

$$y_i = \sum_{j=1}^n a_j x_{(j)}. \quad (61)$$

The moving average, median, r th ranked-order, α -trimmed mean, and midpoint filters are special cases of (61) if the coefficients a_j , $j = 1, \dots, n$, are chosen appropriately. The

Table 3 Optimal Coefficients of an L Filter Having $n = 9$ Coefficients for Various Input Distributions

Distributions	a_1, a_9	a_2, a_8	a_3, a_7	a_4, a_6	a_5
Uniform	0.5	0.0	0.0	0.0	0.0
Normal	0.11	0.11	0.11	0.11	0.11
Laplacian	-0.01899	0.02904	0.06965	0.23795	0.36469

following choices of the coefficients give the median filter and the midpoint filter respectively:

$$a_j = \begin{cases} 1 & j = \nu + 1 \\ 0 & j \neq \nu + 1 \end{cases} \quad (62)$$

$$a_j = \begin{cases} 1/2 & j = 1, n \\ 0 & j \neq 1, n. \end{cases} \quad (63)$$

L filters are based on the theory of robust L estimators, which form one of the three major families of robust estimators [7], [8]. They have been proposed in the context of digital signal filtering only recently [114]. The filter coefficients a_j , $j = 1, \dots, n$, can be chosen to satisfy an optimality criterion that is related to the probability distribution of the input noise. Let us consider the additive noise model $y_i = s_i + n_i$. The filter coefficients can be chosen in such a way that the error norm is minimized:

$$MSE = E[(s_i - y_i)^2] = E[(\sum_{j=1}^n a_j x_{(j)} - s_i)^2] = \mathbf{a}^T \mathbf{R} \mathbf{a}, \quad (64)$$

where \mathbf{R} is the $n \times n$ correlation matrix of the vector of the ordered noise variables $\mathbf{n} = [n_{(1)}, \dots, n_{(n)}]^T$. \mathbf{a} is the coefficient vector $[a_1, \dots, a_n]^T$. A *location invariance* constraint must be imposed on the estimator (61). This constraint means that a shift in the signal samples from x_i to $x_i + c$ leads to a shift in the filter output from y_i to $y_i + c$. The filter coefficients must satisfy the following relation in this case:

$$\sum_{j=1}^n a_j = \mathbf{a}^T \mathbf{e} = 1, \quad (65)$$

where $\mathbf{e} = [1, \dots, 1]^T$. The minimization of the Lagrangian function $F(\lambda, \mathbf{a})$ of the mean square error and the constraint function

$$F(\lambda, \mathbf{a}) = \mathbf{a}^T \mathbf{R} \mathbf{a} + \lambda(\mathbf{a}^T \mathbf{e} - 1) \quad (66)$$

gives the following coefficient vector:

$$\mathbf{a} = \frac{\mathbf{R}^{-1} \mathbf{e}}{\mathbf{e}^T \mathbf{R}^{-1} \mathbf{e}}. \quad (67)$$

It can be shown, by using the Cauchy-Schwartz inequality, that the corresponding mean square error is always less than or equal to that produced by the arithmetic mean [115]. Therefore, the L filter will never perform worse than the moving average filter, in the mean square sense. The filter coefficients depend entirely on the correlation matrix \mathbf{R} , i.e., on the input noise probability function. If the

input noise distribution is Gaussian, the filter coefficients are given by $a_j = 1/n$, $j = 1, \dots, n$. As expected, the optimal L filter for the Gaussian noise is the moving average filter. The optimal L filter for the Laplacian distribution is close to the median, as can be seen in Table 3. The optimal L filter for the uniform distribution is the midpoint. The above mentioned analysis of L filters is basically applied to constant signals corrupted by additive noise. It can be extended to the case of arbitrary signals corrupted by additive iid noise. Structural constraints can be incorporated in the optimization function in order to design filters that are sensitive to local signal structures [116]. Similar techniques can be applied to the optimal design of L filters for dependent noise, e.g. Markov noise [117]. The ability of the L filter to have optimal coefficients for a variety of input distributions makes it suitable for a large number of applications. A further advantage of the L filter over the median is that it has no streaking effects, provided that its coefficients are not similar to that of the median filter. Its disadvantage over both the median and the moving average filter is that it has greater computational complexity, because its calculation requires additions, multiplications, and comparisons. The L filter has poor robustness properties if all its coefficients are nonzero [118], [119]. Its robustness to impulses is improved by trimming out small and large data before summation [120]. The deterministic properties of the L filters and their relation to linear filters are discussed in [123]. L filters have successfully been applied to ultrasonic image processing, where speckle noise has a Rayleigh distribution [121], [122]. It has been proved that the large observations in the filter window are weighted heavily. Its performance is better than that of other filters (e.g. median filter, moving average filter) in speckle noise removal. Generalizations of L filters can be found in [124].

One of the main disadvantages of L filters and of other filters based on order statistics is that ordering destroys time or neighborhood information. Therefore, their performance deteriorates when the filter length, n , increases beyond a certain length [125]. This poor performance is particularly evident in the case of nonstationary signals. A modification of L filters has been proposed which takes into account time information. These filters are called LI filters [125] or C filters [126]. Their definition is the following:

$$y_i = \sum_{j=1}^n a(R(i - j + \nu + 1), j) x_{i-j+\nu+1}. \quad (68)$$

The filter coefficients $a(R(i - j + \nu + 1), j)$, $j = 1, \dots, n$

depend on the position $i-j+\nu+1$ of the sample $x_{i-j+\nu+1}$ as well as on the rank $R(i-j+\nu+1)$ of this sample in the sample set $x_{i-\nu}, \dots, x_{i+\nu}$. Thus C filters can be considered to be time-varying L filters. Therefore, they can be easily adapted to time-varying environments [125].

E. R Filters

These are based on another large class of robust estimators, the so-called R estimators [7], [8], [127]. The most important R filter is the *Wilcoxon filter* [128]–[131]:

$$y_i = \text{med}\left\{\frac{x_{(j)} + x_{(k)}}{2}, 1 \leq j \leq k \leq n\right\}, \quad (69)$$

where $x_{(j)}, x_{(k)}$, $j, k = 1, \dots, n$, are the order statistics of the samples $x_{i-\nu}, \dots, x_{i+\nu}$. An alternative definition of the Wilcoxon filter is the following:

$$y_i = \text{med}\left\{\frac{x_j + x_k}{2}, i-\nu \leq j \leq i+\nu, i-\nu \leq k \leq i+\nu\right\}. \quad (70)$$

Wilcoxon filters have been proved to be effective in the removal of additive Gaussian noise. However, they do not preserve edges well, because the summation in (69) uses every possible pair for $1 \leq j \leq k \leq n$. If the summation in (69) is restricted to a maximum distance $j-k \leq D$, the *modified Wilcoxon filter* results [130]:

$$y_i = \text{med}\left\{\frac{x_{(j)} + x_{(k)}}{2}, 1 \leq j \leq k \leq n, k-j < D\right\}. \quad (71)$$

By letting D range from 1 to n , the modified Wilcoxon filter ranges from the median ($D = 1$) to the Wilcoxon filter ($D = n$). The modified Wilcoxon filter has better edge preservation properties than the standard Wilcoxon filter, but it is still worse than the median filter in this respect. A further disadvantage of the Wilcoxon filter is its computational complexity. It requires the ordering of n numbers, $n(n+1)/2$ additions, and the ordering of $n(n+1)/2$ numbers. Algorithms for fast Wilcoxon filter calculation are proposed in [132]. Other modifications of the Wilcoxon filter can be found in [133] and [134].

F. M Filters M filters are not directly related to order statistics. However, they are included here because they come from one of the best known families of robust estimators: the *M estimators*. Their definition is the following [7], [8]:

$$\sum_{j=i-\nu}^{i+\nu} \psi(x_j - y_i) = 0. \quad (72)$$

Here $\psi(x)$ is generally an odd, continuous, and sign-preserving function; y_i is the output; and x_j , $j = i-\nu, \dots, i+\nu$, are the input samples in the filter window. A special case of the M filters is the *maximum likelihood filters*, having

$$\psi(x) = \frac{f'(x)}{f(x)}, \quad (73)$$

where $f(x)$ is the probability distribution of the data x_i . The moving average filter is a maximum likelihood filter

for Gaussian distributed data:

$$\sum_{j=i-\nu}^{i+\nu} (x_j - y_i) = 0. \quad (74)$$

The maximum likelihood filter of Laplacian distributed iid data is the median filter, satisfying

$$\sum_{j=i-\nu}^{i+\nu} \text{sign}(x_j - y_i) = 0. \quad (75)$$

Therefore, both the moving average and the median are special cases of M filters. Let us suppose that the model data cdf is $F(x)$ and that the data contain $\varepsilon\%$ outliers having an unknown symmetric distribution $H(x)$. In this case, the data distribution $G(x)$ lies in the neighborhood P_ε of the distribution $F(x)$:

$$P_\varepsilon = \{(1-\varepsilon)F + \varepsilon H; H \text{ is a symmetric distribution.}\} \quad (76)$$

The M estimator (called the Huber estimator) which minimizes the asymptotic variance $V(\psi, G)$ over P_ε and the corresponding M filter are defined by using the following function in (72) [7], [8]:

$$\psi(x) = x \min\left(1, \frac{b}{|x|}\right). \quad (77)$$

For Gaussian distributed data, the constant b is given by an implicit equation:

$$2\Phi(b) - 1 + 2\phi(b)/b = 1/(1-\varepsilon) \quad (78)$$

where $\Phi(x), \phi(x)$ denote the cdf and the pdf of the unit Gaussian distribution respectively [7]. The function $\psi(x)$ is plotted in Fig. 14. The Huber estimator and the corresponding M filter, called the standard M filter (STM) [112], are B robust and they have breakdown point $\varepsilon^* = 1/2$; i.e. they can reject up to 50% outliers [7]. The Huber estimator tends to the median or to the arithmetic mean when b tends to zero or to infinity respectively. Therefore, the STM filter is another compromise between the median filter and the moving average filter. If all data x_j , $j = i-\nu, \dots, i+\nu$, are far away from their median, all differences $x_j - y_i$ lie in the saturated region of $\psi(x)$; therefore, the Huber estimator coincides with the median. On the other hand, if all data are close to their median and to each other, the differences $x_j - y_i$ lie in the linear region of $\psi(x)$. In this case, the STM filter coincides with the moving average filter. In the homogeneous image regions, the STM filter behaves like a moving average filter. Its performance is similar to that of the median close to image edges. Impulsive noise can be effectively removed by the STM filter, owing to its robustness properties. A detailed analysis of the performance of the STM filter can be found in [135]. A modification of the M filter, called the W filter, is described in [136]. Despite its advantages, the STM filter is not very popular because of its implicit definition, which requires iterative techniques for the calculation of its output. However, it has been observed that Newton iteration for the solution of (72) requires only five iterations with absolute error less than 0.01 [112].

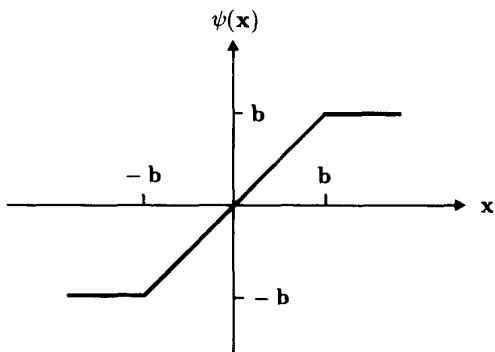


Fig. 14. The function which is used in the definition of the Huber estimator and the corresponding M filter.

G. General Nonlinear Filter Structure

As described in the introduction, several nonlinear filter classes have been used in image processing, e.g. filters based on order statistics, homomorphic filters [137], polynomial filters [138]–[146], and nonlinear mean filters [109], [147]. Since nonlinear filters come from different theories, their structure and properties vary widely. Therefore, a unifying nonlinear filter structure which encompasses several classes of nonlinear filters as special cases is very important. Such a filter structure is described in Fig. 15. It consists of two pointwise-nonlinear functions, $g(x)$, $f(x)$, a sorting network and two linear filters, L_1 and L_2 . Filters based on order statistics, homomorphic filters, nonlinear mean filters, and morphological filters are special cases of the general nonlinear filter structure if the functions $g(x)$ and $f(x)$ and the linear filters are chosen appropriately [148]. A special case of the general nonlinear filter structure, called the nonlinear order statistic filter (NLOS), shown in Fig. 16, can be used for signal-dependent noise filtering described by

$$y = t(s) + r(s)n_1 + n_2. \quad (79)$$

The nonlinear function $g(x)$ is chosen in such a way that signal-dependent noise is transformed to additive white noise:

$$g^{(1)}[t(s)] = \frac{1}{r(s)}. \quad (80)$$

The coefficients a_i , $i = 1, \dots, n$, are given by (67). The nonlinear function $f(x)$ is used to restore the original dynamic range of the image by choosing

$$f(s) = g^{-1}[t(s)]. \quad (81)$$

The NLOS filter is a nonlinear filtering module that can perform as a median filter, L filter, α -trimmed mean filter, r th rank order filter, moving average filter, midpoint filter, erosion or dilation filter, range edge detector, or dispersion edge detector [150]. Therefore, it can be used as a versatile building block for nonlinear filtering. Such a 3×3 building block is shown in Fig. 17. This module

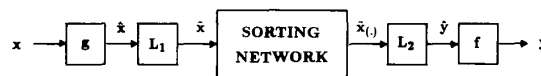


Fig. 15. General nonlinear filter structure.

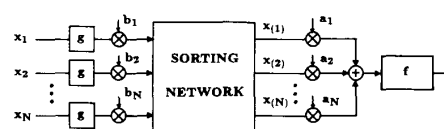


Fig. 16. Nonlinear order statistics filter structure.

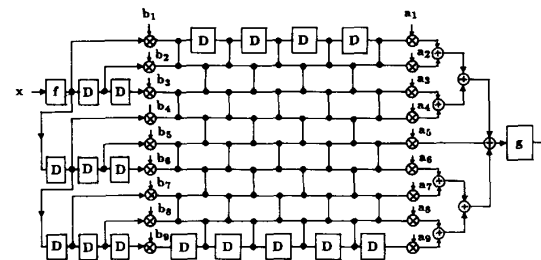


Fig. 17. A 3×3 general nonlinear filtering module. D denotes delay. Vertical bars denote comparators.

has a very regular and parallel structure and it uses small interconnection paths. Therefore it is very suitable for VLSI implementation. Its hardware requirements are equal to the requirements of a moving average filter and a median filter. Its throughput delay is the sum of the throughput delay of a median and of a moving average filter. This is the price paid for its versatility. An analysis of the computational complexity of the nonlinear filtering module can be found in [148].

H. Multichannel Median Filters Multichannel filtering is very important in color image processing, in multiband image processing for remote sensing applications, in velocity field filtering, and in complex-valued signal filtering. Multichannel color image filtering has attracted particular attention in recent years, because color images are essentially three-channel two-dimensional signals [151]–[155]. Color images can be described as vector images. There has been a particular interest in extending nonlinear black and white image processing techniques based on order statistics to color images [156]–[162]. Single-channel order statistic filtering uses the notion of data ordering, which is very natural in this case. This notion cannot be extended in a straightforward way in the case of multivariate data. An excellent treatment of multivariate data ordering can be found in [164]. It is shown that there are several ways to order multivariate data. There is no unambiguous, universally agreeable total ordering of the n p -variate samples $\mathbf{x}_1, \dots, \mathbf{x}_n$ $\mathbf{x}_i = [x_{1i}, x_{2i}, \dots, x_{pi}]^T$, $i = 1, \dots, n$. The following so-called subordering principles are discussed in [164]: *marginal ordering*, *reduced (aggregate) ordering*, *partial ordering*, and *conditional (sequential) ordering*.

In marginal ordering, the multivariate samples are ordered along each of the p dimensions:

$$\begin{aligned} x_{1(1)} &\leq x_{1(2)} \leq \dots \leq x_{1(n)} \\ x_{2(1)} &\leq x_{2(2)} \leq \dots \leq x_{2(n)} \\ &\vdots \\ x_{p(1)} &\leq x_{p(2)} \leq \dots \leq x_{p(n)} \end{aligned} \quad (82)$$

i.e., ordering is performed in each channel of the multichannel signal. The i th marginal order statistic is the vector $\mathbf{x}_{(i)} = [x_{1(i)}, x_{2(i)}, \dots, x_{p(i)}]^T$. The cdf and the pdf of the marginal order statistics can be found in [156] and [165]–[167]. The extension of the notion of influence function to the marginal order statistics can be found in [7] and [156]. Marginal order statistics filters have been applied in color image processing [156], [157]. Their behavior in impulse noise removal is studied in [158]. The reduced ordering (R ordering) is based on the generalized distance:

$$d = (\mathbf{x} - \mathbf{a})^T \Gamma^{-1} (\mathbf{x} - \mathbf{a}) \quad (83)$$

of a sample \mathbf{x} from a point \mathbf{a} which may be either the origin or the sample arithmetic mean $\bar{\mathbf{x}}$ or the marginal median $\mathbf{x}_{(\nu+1)}$. Γ may be the identity matrix, the dispersion matrix \mathbf{C} , or the sample dispersion matrix \mathbf{S} . The various data \mathbf{x}_i are ordered according to their distances d_i from \mathbf{a} , as shown in Fig. 18. Thus multivariate ordering is reduced to one-dimensional ordering. Reduced ordering has been used in color image filtering [168] with very promising results. It has been observed that the Mahalanobis distance (83) gives better results than the Euclidean distance ($\Gamma = \mathbf{I}$).

Partial ordering (P ordering) is based on the notion of the convex hull of the points $\mathbf{x}_1, \dots, \mathbf{x}_n$, which is the minimum convex set which encloses all n samples [5], [169]. The conditional ordering (C ordering) is conducted on one of the marginal sets of observations conditional on ordering within the data in terms of other marginal sets of observations [164].

The definitions of L estimators can be easily extended to the p -dimensional case by using marginal order statistics. The following estimator will be called the p -dimensional marginal L -estimator [156]:

$$\mathbf{T}_n = \sum_{i_1=1}^n \dots \sum_{i_p=1}^n \mathbf{A}_{i_1, \dots, i_p} \mathbf{x}_{(i_1, \dots, i_p)}, \quad (84)$$

where $\mathbf{x}_{(i_1, \dots, i_p)} = [x_{1(i_1)}, \dots, x_{p(i_p)}]^T$ are the marginal order statistics and $\mathbf{A}_{i_1, \dots, i_p}$ are $p \times p$ matrices. The performance of the marginal L estimator depends on the choice of these matrices. The marginal median, maximum, and minimum are special cases of (84) for appropriate choices of the matrices $\mathbf{A}_{i_1, \dots, i_p}$. The p -dimensional marginal α -trimmed mean [170]:

$$\mathbf{T}_n = \left[\begin{array}{c} \frac{1}{n(1-2\alpha_1)} \sum_{i=\alpha_1 n+1}^{n-\alpha_1 n} x_{1(i)} \\ \dots \\ \frac{1}{n(1-2\alpha_p)} \sum_{i=\alpha_p n+1}^{n-\alpha_p n} x_{p(i)} \end{array} \right] \quad (85)$$

is another special case of the p -dimensional L estimator.

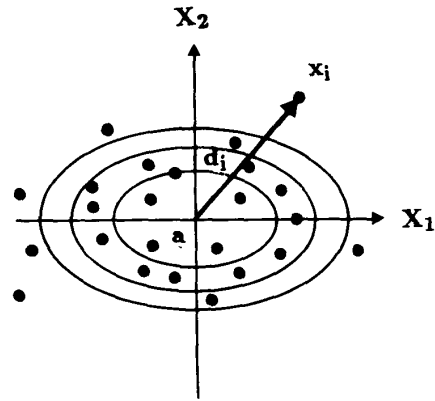


Fig. 18. R ordering of multivariate data.

L estimators can be used in multichannel image filtering in the case of additive white noise:

$$\mathbf{x} = \mathbf{s} + \mathbf{n}, \quad (86)$$

where \mathbf{s} is the desired signal, \mathbf{n} is a white noise process, and \mathbf{x} is the corrupted signal. The marginal median, which is a special case of marginal L estimators, has been studied in [156]. It has been shown that it behaves well in the case of additive white long-tailed noise and in the case of impulsive noise. Multichannel L filters have been developed for various multichannel distributions in [159] and have been compared to single channel L filters and to marginal filters (e.g. marginal median filter and marginal moving average filter). It has been found that multichannel L filters outperform single-channel filters if the signal channels are correlated to each other. Another definition of the multichannel median, called vector median, has been proposed for color image filtering [160], [162], [163]. It is the vector \mathbf{x}_{med} that minimizes the L_1 error norm:

$$\sum_{i=1}^n |\mathbf{x}_i - \mathbf{x}_{med}| \rightarrow \min. \quad (87)$$

This definition of the multichannel median is a direct extension of the corresponding single-channel median definition. Generally, it gives results different from those of the marginal median. The vector \mathbf{x}_{med} may or may not be one of the data \mathbf{x}_i , $i = 1, \dots, n$. The implicit definition (87) of the vector median makes its application difficult in color image filtering because the use of an iterative optimization algorithm is required to produce each output image pixel. The computational load is reduced if the median \mathbf{x}_{med} is forced to belong to the set \mathbf{x}_i , $i = 1, \dots, n$. In this case, the L_1 norm (87) is computed for every $\mathbf{x}_{med} = \mathbf{x}_i$, $i = 1, \dots, n$. The median is the one that minimizes [162]:

$$\begin{aligned} \sum_{i=1}^n |\mathbf{x}_i - \mathbf{x}_{med}| &\leq \sum_{i=1}^n |\mathbf{x}_i - \mathbf{x}_j| \quad j = 1, \dots, n, \\ \mathbf{x}_{med} &\in \{\mathbf{x}_1, \dots, \mathbf{x}_n\}. \end{aligned} \quad (88)$$

Other distance measures (e.g. Mahalanobis distance) can

be used in (88) instead of the L_1 norm. The statistical and deterministic properties of the vector median filters are described in [162]. Application of vector median filters in velocity filtering and in complex narrow-band signals can be found in [162] and [161] respectively. Median hybrid filters have also been extended for vector signals [160]–[162].

A comparison of the marginal median filters, vector median filters, in color image filtering is included in [158]. Marginal median filters and vector median filters have similar performance. Both of them are outperformed by the multichannel α -trimmed mean filters and the multichannel MTM and DW-MTM filters. Multichannel DW-MTM filters clearly have the best performance in all simulations done both for one-dimensional signals and for color images.

IV. ADAPTIVE ORDER STATISTIC FILTERS

The nonlinear filters described in the previous chapters are usually optimized for a specific type of noise and sometimes for a specific type of signal. However, this is not usually the case in many nonlinear filtering applications, especially in image processing. Images can be modeled as two-dimensional stochastic processes, whose statistics vary in the various image regions. Images are nonstationary processes. Furthermore, noise statistics, e.g., the noise standard deviation and even the noise probability density function, vary from application to application. Sometimes, noise characteristics vary in the same application from one image to the next. This is the case in the channel noise encountered in image transmission and the atmospheric noise (e.g., the cloud noise) in satellite images. In such environments, nonadaptive filters cannot perform well, because their characteristics depend on the noise and signal characteristics, which are unknown. Therefore, adaptive filters are the natural choice in such cases.

Adaptive filter performance depends on the accuracy of the estimation of certain signal and noise statistics, namely the signal mean and standard deviation and the noise standard deviation. The estimation is usually local; i.e., relatively small windows are used to obtain the signal and noise characteristics. Such an adaptive filter can be employed for additive white noise [171], [172]:

$$x_{ij} = s_{ij} + n_{ij} \quad (89)$$

The linear minimal mean square error estimate (MMSE) of s_{ij} is given by the following formula:

$$\hat{s}_{ij} = \left(1 - \frac{\sigma_n^2}{\sigma_x^2}\right)x_{ij} + \frac{\sigma_n^2}{\sigma_x^2}\hat{m}_x, \quad (90)$$

where σ_n , σ_x , and \hat{m}_x are the local estimates of the noise standard deviation, the signal standard deviation, and the signal mean, respectively. At homogeneous image regions, the noise standard deviation is approximately equal to the signal standard deviation. In these regions, the adaptive MMSE filter (90) is reduced to the local estimate of the signal mean $\hat{s}_{ij} \approx \hat{m}_x$. At edge regions, the signal standard

deviation is much greater than the noise signal deviation ($\sigma_n \ll \sigma_x$). In these regions, no filtering is performed at all ($\hat{s}_{ij} = x_{ij}$). Thus, the adaptive MMSE filter preserves edges, although it does not filter the noise in the edge regions. The performance of the adaptive MMSE filter depends on the choice of the local measures of signal mean and standard deviation and of the noise standard deviation. The local arithmetic mean and sample standard deviation have been used in [171], [172] for the estimation of the signal mean and standard deviation, respectively. The local median and the midpoint were proposed in [173] and [174] for the estimation of the signal mean. Another approach that uses adaptive data trimming was proposed in [175]. The resulting filter is called the adaptive double window modified trimmed mean (DW MTM) filter and it is used in signal-dependent noise filtering.

Another reason for using adaptive filters is edge preservation. Certain filters, e.g., the moving average filter, perform well in homogeneous image regions but fail close to edges. The opposite is true for other filters, e.g., the median filter. A combined filter which performs differently on image edges than in image plateaus can be used. Such filters are also called decision directed filters because they employ an edge detector to decide if an edge is present or not. Order statistics are efficient tools in edge detection. The *range* W and the *quasi range* $W_{(i)}$ can be used as edge detectors [176]:

$$W = x_{(n)} - x_{(1)} \quad (91)$$

$$W_{(i)} = x_{(n+1-i)} - x_{(i)}, \quad 2 \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor. \quad (92)$$

The range edge detector can be calculated easily. However, it has poor robustness to impulses. The *dispersion edge detector* [176]:

$$J_i = W + W_{(1)} + W_{(2)} + \cdots + W_{(i)} \quad (93)$$

has better robustness properties than the range edge detector. The comparison of the median filter output at neighboring windows can also be used as an edge detector [177]. If the difference between the two outputs is larger than a threshold, an edge is declared. Another approach to edge detection is to use *rank tests* [178]. The *Wilcoxon test* and the *median test* have been proposed in [179].

A decision directed filter is shown in Fig. 19. It consists of two L filters [150]. One is performing as edge detector (e.g. range edge detector) and the other is performing as a usual L filter. The output of the edge detector controls the coefficients of the second L filter. If no edge is detected, the L filter operates as a moving average filter. If an edge is detected, the L filter operates as a median filter.

Decision-directed filters can take into account both edge information and impulsive noise information. Impulses, when detected, are not taken into account in the estimation of local mean and standard deviation. Furthermore, when an edge is detected, the window of the filter can become smaller so that edge blurring is minimized. Such an impulse

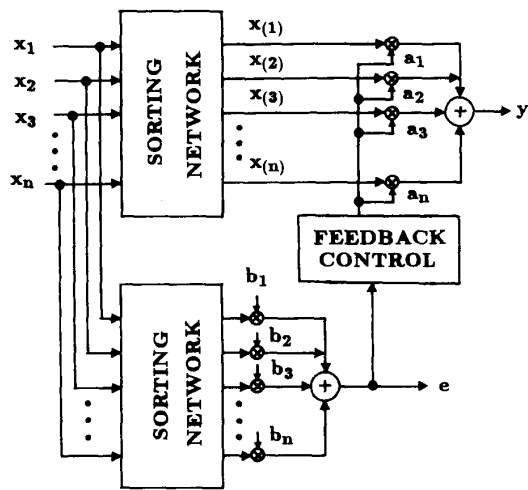


Fig. 19. Decision-directed filter based on two L filters.

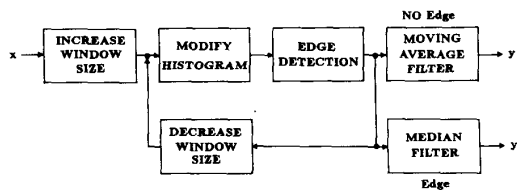


Fig. 20. Adaptive window edge detection (AWED) filter.

sensitive filter, called the adaptive window edge detection (AWED) filter, is shown in Fig. 20 [180]. The filter initially starts with a 7×7 or a 5×5 window. The local image histogram in the filter window is examined. If impulses are detected, they are rejected and the local image standard deviation calculation is based on the rest of the pixels in the window. If the local standard deviation is low enough, a homogeneous image region is assumed and the moving average filter is used. If the local standard deviation is large (above a certain threshold) an edge is declared. If the window size is 3×3 , the median filter is used for image filtering. If the window size is greater than 3×3 , it is reduced and the whole procedure is repeated. The window size is increased at each pixel if no edge has been detected.

An adaptive version of the α -trimmed mean filter has been proposed in [181]. Order statistics are used as estimators of the data distribution tail [182]. Based on this estimator, a decision is made on whether to use the mid-point filter, the complementary α -trimmed mean filter, the moving average filter, the α -trimmed mean filter, or the median filter. An adaptive median filter based on decisions of hypothesis tests is presented in [183].

Adaptive versions of morphological filters have also been reported recently [184], [185]. The operational filter windows can adapt their shapes according to the local image features.

Another approach related to decision directed filtering is the *two-component model* filtering. An image x can be

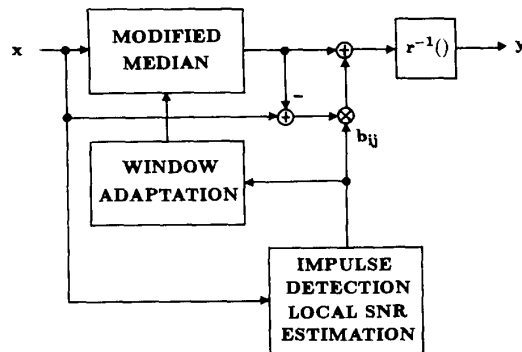


Fig. 21. Signal-adaptive median (SAM) filter structure.

considered to consist of two parts: a low-frequency part x_L and a high-frequency part x_H :

$$x = x_L + x_H. \quad (94)$$

The low-frequency part is dominant in homogeneous image regions, whereas the high-frequency part is dominant at edge regions. The two-component image model allows different treatment of its components. Therefore, it can be used for adaptive image filtering and enhancement, provided that the two components can be separated. A low-pass and a high-pass filter can be used for the separation of the two components. In most cases, the moving average filter or the median filter are used as estimators \hat{m}_x of the low-frequency component [187], whereas the high-frequency component is given by

$$x_{Hij} = x_{ij} - \hat{m}_x. \quad (95)$$

An adaptive filter based on the two-component image model is shown in Fig. 21. It is called a signal-adaptive median (SAM) filter [188]. Its output signal is given by

$$y_{1ij} = \hat{m}_x + b_{ij}(x_{ij} - \hat{m}_x) \quad (96)$$

$$y_{ij} = r^{-1}(y_{1ij}). \quad (97)$$

The performance of the adaptive filter ((96) and (97)) depends on the choice of the coefficient b_{ij} , which can be done in an optimal way for various noise types (e.g. additive and multiplicative noise) [188]. This coefficient also detects edge information. Thus, b_{ij} can be used for the adaptation of the filter window. We can start filtering by using initial window size 5×5 or 7×7 . If the coefficient b_{ij} becomes greater than an appropriate threshold, b_t , close to image edges, the window size is decreased until the coefficient becomes less than the threshold or until the window reaches the size 3×3 . Otherwise, the window size is increased to its maximum size. If impulsive noise is present, the impulses can be detected and removed from the filter window. The local median of the remaining pixels can be used as an estimate of the signal mean. The SAM filter has excellent performance in noise filtering and in edge and image detail preservation, as can be seen in Fig. 22 [180].

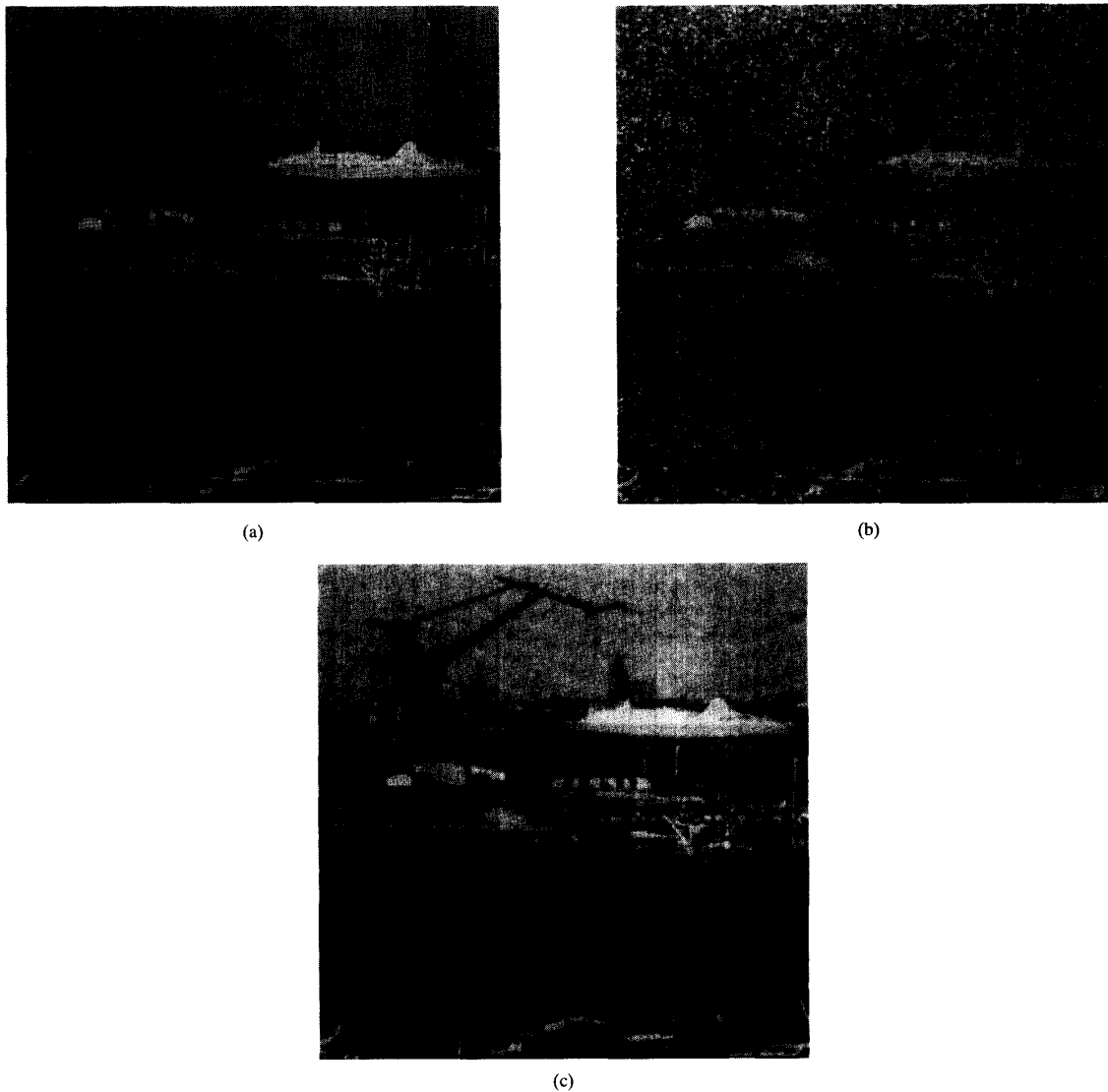


Fig. 22. (a) Original image. (b) Image corrupted by mixed impulsive and additive Gaussian noise. (c) Output of the SAM filter.

Several classes of adaptive nonlinear filters have been presented in this paper so far. Their comparative study is rather difficult. Such comparisons and quality/complexity trade-offs are presented in [189] and [5]. Most adaptive filters presented in this section are based on rather ad hoc or heuristic techniques to derive the filter parameters and, sometimes, also the filter structure itself. In contrast, a rigorous mathematical theory has been developed for linear adaptive filters [190], [191]. Therefore, an effort has been made recently to extend this mathematical approach to nonlinear filters based on order statistics [192]–[203].

Let $d(i)$ be the uncorrupted zero-mean signal, which will also be used as a reference signal for the derivation of the adaptive L filter. The observed signal $x(i)$ is given by

$$x(i) = d(i) + n(i). \quad (98)$$

The adaptive L filter is defined as follows:

$$\hat{d}(i) = y(i) = \sum_{j=1}^n a_j(i)x_{(j)}(i) = \mathbf{a}^T(i)\mathbf{x}(i). \quad (99)$$

The coefficient vector $\mathbf{a}(i) = [a_1(i), \dots, a_n(i)]^T$ must be updated at each step i in such a way that the mean square error J is minimized:

$$J = E[e(i)^2] = E\left[\left[d(i) - \sum_{j=1}^n a_j(i)x_{(j)}(i)\right]^2\right]. \quad (100)$$

It can be easily proved that the filter coefficients must be adapted as follows:

$$\mathbf{a}(i+1) = \mathbf{a}(i) + \mu e(i)\mathbf{x}(i), \quad (101)$$

where μ denotes the step size. Algorithm (101) is equivalent to the *least mean square (LMS) algorithm*, which is very popular in linear adaptive filtering [190], [191]. The only difference is that (101) uses the vector of ordered observations $\mathbf{x}(i) = [x_{(1)}(i), \dots, x_{(n)}(i)]^T$ to update the adaptive L filter coefficients, whereas the LMS algorithm uses the vector $[x(i), \dots, x(i-n+1)]^T$ to update the coefficients of the adaptive FIR filter. Therefore, (101) will be called the LMS L-filter algorithm. The adaptive LMS L filter and its modified schemes are known to adapt well to various noise distributions [196]. They tend to the midpoint, moving average, and median filters for short-tailed, medium-tailed, and long-tailed noise distributions respectively. Location invariance (65) and unbiasedness constraints can be incorporated in LMS adaptation [196]–[198]. RLS algorithms have also been proposed for L-filter adaptation [192], [193], [195]. They have relatively fast convergence compared with LMS L-filter adaptation. The back-propagation algorithm [199], originally used in neural network learning, has also been used in the adaptation of multilayer median filters, median hybrid filters, and L1 filters [200]–[202]. Adaptive L1 filters have been proposed for channel equalization applications [204]. Adaptive recursive L1 filters have also been proposed recently [203]. Finally, median filtering of the error sequence has been used in the LMS adaptation of linear FIR filters in order to improve adaptation performance in the case of impulsive or non-Gaussian noise [205]. Exponential convergence is demonstrated in the mean to the minimum mean squared error solution, which is largely unaffected by impulses [206].

V. ALGORITHMS AND STRUCTURES FOR FILTERS BASED ON ORDER STATISTICS

In recent years, several algorithms and structures have appeared in the literature, especially for median filters. Many such structures intended for VLSI implementation are based on classical *sorting networks* [207]. The most popular one is the *odd/even transportation* structure, which is shown in Fig. 23. It has a very regular parallel structure and short interconnection paths. It has been used as a building block of a median filtering module [208] and for the nonlinear filter module shown in Fig. 17 [148]. The first major disadvantage of this network is that, once it is built to sort n data points, it cannot be extended easily to sort a larger number of data. For example, if a three point sorter is available, seven such sorters are required for the sorting of nine numbers. Thus, the hardware requirements for the extension from three to nine point are excessive. Such a sorting structure was proposed in [209] and is shown in Fig. 24. Another disadvantage is that it requires $n(n-1)/2$ comparators to sort n data points. Other median filter architectures are based on bit-plane median computation by using threshold decomposition techniques [90]–[92]. A review and comparison of the various median filter architectures proposed for VLSI implementation can be found in [210]. With the development of the technology for switched capacitor (SC) filters [211] and charge

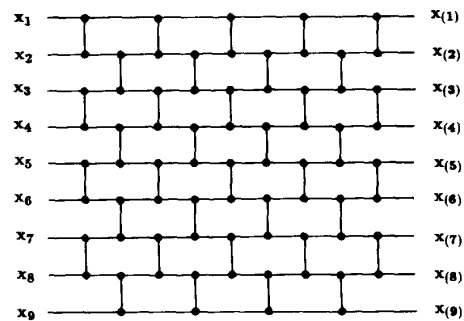


Fig. 23. Odd/even transportation network. The vertical bars denote comparators.

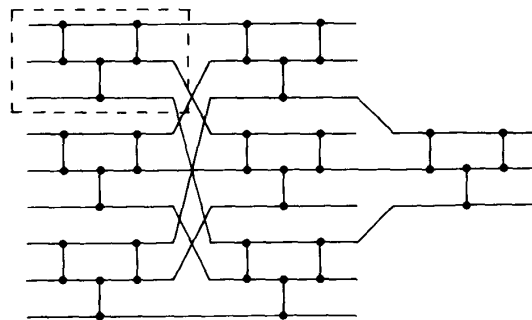


Fig. 24. Sorting network for nine numbers whose building block is a sorting network for three numbers (indicated by a block at the upper left corner). The vertical bars denote comparators.

coupled device (CCD) [212] filters and delay lines, analog techniques have been proposed for the implementation of certain nonlinear filters. The analog implementation of a median filter is described in [213]. Also analog/digital and switched capacitor implementations of the median filters and of the median hybrid filters have been proposed recently [214]–[217].

Several algorithms have been reported in the literature for the calculation of the median and for data sorting [207]. BUBBLE-SORT, QUICK-SORT, and MERGE-SORT are among the most popular sorting algorithms. Since sorting algorithms do not require many more comparisons than the algorithms for the calculation of the median [207], sorting algorithms are used also in median filtering. One of the fastest algorithms is QUICK-SORT. It requires $O(n \ln n)$ comparisons in the average and $O(n^2)$ comparisons in the worst case for the sorting of n numbers [218]. However, QUICK-SORT is not the fastest algorithm to be used in median filtering. Median filters require the calculation of running medians. From one output pixel to the next, the $n \times n$ filter window moves by one column. Thus, n input pixels are removed and n new pixels enter the filter window. The remaining $n^2 - 2n$ pixels remain unchanged. If this fact is taken into account, very fast running median algorithms can be constructed. Huang *et al.* [219] have implemented such an algorithm. It is based on the gray level histogram of the

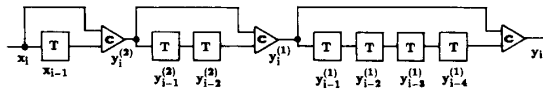


Fig. 25. A structure for the implementation of an eight point max/min filter. T denotes delay and C denotes a comparator.

n^2 pixels in the filter window. This histogram is stored and updated as the window moves. The number of comparisons per output point required by this algorithm is $O(n)$, which is far less than the number, $O(n^2 \ln n)$, required by the QUICK-SORT algorithm for $n \times n$ numbers. Another fast running median algorithm is presented in [220]. Similar fast running algorithms for max/min calculation and sorting and for recursive median filtering are given in [221] and [222]. The number of comparisons per output point required by the fast running algorithm for an $n \times n$ max/min filter is only 6. This number is much less than the $n^2 - 1$ comparisons required by the standard max/min calculation algorithms [218]. Furthermore, the number of comparisons required does not increase with the filter dimension n . The number of comparisons required by the fast running sorting algorithm for an $n \times n$ window is of the order $O(n \log_2 n)$, which is very close to that of the Huang's algorithm and is much lower than that of the QUICK-SORT algorithm [221]. Fast algorithms for separable median filtering can be found in [223]. A comparison of the computation speeds of various fast median algorithms can be found in [224] and [225]. The running max/min calculation and sorting algorithms can also produce very efficient processor structures for max/min and sorting calculations. Such a structure for the calculation of an eight-point max/min filter is shown in Fig. 25 [221]. Only three comparisons are required per output point. Finally, fast median filtering algorithms have been proposed which are based on bit-plane operations [226]–[231]. The performance of these algorithms depends on the number of bits used for the representation of a single pixel.

VI. COMPARISONS AND CONCLUSIONS

Robust estimation theory has provided some excellent tools for nonlinear filtering. Order statistics is such a valuable tool. Its use in digital filtering, especially for image processing applications, has produced an entire family of nonlinear filters. The best known and most widely used order statistic filter is the median filter. Its widespread use is based on its simplicity, its calculation speed, and its excellent edge preservation and impulse removal properties. The median filter has been fairly well studied from a theoretical point of view. Its disadvantages, mainly the streaking effect, the destruction of fine image details, and its relatively poor performance in additive Gaussian noise filtering, have led to the development of other order statistic filters. Several of them are essentially a compromise between median and moving average filters. Such filters are the α -trimmed filter, the L filter, and the STM filter. Their characteristics can be tailored to the noise probability distribution. Therefore,

they are useful in a variety of applications. Other filters try to incorporate spatial information in the filter structure, aiming at the preservation of image details. Such filters are the max/median filter and the median hybrid filters. Finally other filters, e.g. the STM filters and the Wilcoxon filters, come from widely used families of robust estimators. Their main disadvantage is their high computational complexity in comparison with the corresponding nonlinear filters based on order statistics.

The multitude of filters poses some difficulties to the design/applications engineer. Most of the filters have their advantages and disadvantages. Therefore, it is not clear which type of filter is suitable for a specific application. This problem arises from the fact that most filters are designed to perform well in the presence of certain types of noise. Usually, their performance deteriorates rapidly in the presence of different types of noise. The most commonly used figures of merit for nonlinear filter performance in digital image processing are the following:

1. noise filtering characteristics for different types of noise;
2. edge preservation;
3. fine detail preservation;
4. unbiasedness;
5. computational complexity.

The noise filtering properties of a filter are usually measured by its output variance or by the rate of success in the impulsive noise removal. Other performance measures are the normalized mean square error (NMSE), the perceptual mean square error (PMSE), and the peak-to-peak SNR (PSNR) [5], [232]–[234]. The mean absolute error (MAE) has been decomposed into impulse error and edge error in order to find perceptually significant error measures for order statistics filtering [235]. The main disadvantage of all these measures (except perhaps the rate of success in impulsive noise removal) is that they are relatively poorly correlated to subjective human criteria. Human vision is rather complicated and its properties cannot be described by a single performance formula.

Edge preservation is an important property of an image filter and it refers to its capability to preserve edges. The fine detail preservation properties refer to its ability to preserve image lines, sharp corners, and other fine image details. The bias of a filter refers to its directional or illumination bias. It is known that certain filters tend to enhance image along certain directions (e.g., along the horizontal and the vertical direction for separable filters). Other filters (e.g., the erosion or dilation filters) tend to enhance image regions having certain illumination characteristics (e.g., low or high illumination). All figures of merit (2–4) are rather qualitative and they have not been described yet by quantitative criteria. Therefore, the performance of the filter according to the figures of merit (2–4) is relatively subjective. Most researchers demonstrate the performance of their filters by using digital noise generators to corrupt an image and by comparing the filtered images with the original images.

Table 4 Overview of the Performance of Various Nonlinear Filters

Filter	Figures of Merit											
	a	b	c	d	e	f	g	h	i	j	k	l
Moving average	1	2	0	0	0	0	0	0	0	0	2	1
Median	0	1	2	2	2	2	0	0	2	0	2	1
Separable median	0	1	2	2	2	2	0	0	2	0	2	2
Recursive median	0	1	2	2	2	2	0	0	2	0	2	1
Max median	0	1	1	1	2	1	0	0	2	1	0	1
Multistage median	0	1	2	2	2	2	0	0	2	2	2	1
Median hybrid	1	1	1	1	1	1	0	0	2	2	2	2
Low ranked order	1	1	0	2	0	0	0	0	2	1	0	1
High ranked order	1	1	0	0	2	0	0	0	2	1	0	1
L_p -mean	1	1	0	0	2	0	0	0	1	1	1	0
CH_p -mean	1	1	0	0	2	0	0	0	2	1	0	0
L_p -mean	1	1	0	2	0	0	0	0	1	1	1	0
CH_p -mean	1	1	0	2	0	0	0	0	2	1	0	0
Harmonic mean	1	1	0	2	0	0	0	0	1	1	1	0
Geometric mean	1	1	1	0	0	0	2	0	0	0	2	0
α -trimmed mean	0	1	2	2	2	2	0	0	2	1	2	0
compl α -trimmed mean	2	1	0	0	0	0	0	0	0	0	2	0
Midpoint	2	1	0	0	0	0	0	0	0	0	2	2
MTM filter	1	1	2	2	2	2	1	1	2	1	2	1
L filter	2	2	2	2	2	2	0	0	2	1	2	0
NLOS filter	2	2	2	2	2	2	2	2	2	1	2	0
Wilcoxon	1	1	1	1	1	1	0	0	0	1	2	0
Modified Wilcoxon	1	1	1	1	1	1	0	0	1	1	2	0
STM filter	1	2	2	2	2	2	0	0	2	1	2	0
Dilation (maximum)	1	1	0	0	2	0	0	0	2	0	0	2
Erosion (minimum)	1	1	0	2	0	0	0	0	2	0	0	2
Closing	1	1	0	0	2	0	0	0	2	0	2	1
Opening	1	1	0	2	0	0	0	0	2	0	2	1
Close-opening	1	1	1	1	2	1	0	0	2	0	2	0
Open-closing	1	1	1	2	1	1	0	0	2	0	2	0

- a. Short-tailed additive white noise
- b. Gaussian additive white noise
- c. Long-tailed additive white noise
- d. Positive impulses
- e. Negative impulses
- f. Salt-pepper noise

- g. Multiplicative noise
- h. Additive signal-dependent noise
- i. Edge preservation
- j. Detail preservation
- k. Bias
- l. Computational complexity

Performance

0: poor

1: average

2: good

Bias

0: strong

1: average

2: low

Computational complexity

0: high

1: average

2: low

The computational complexity of a filter usually refers to the number of algebraic operations (multiplications, comparisons, additions) required per output pixel. In the case of the parallel computation, it refers to the number of hardware resources (adders, multipliers, comparators) required and to the throughput delay per output pixel. Throughput delay is also related to the least possible parallel computation time (called critical time). For completely serial computation on a general-purpose computer, the computational complexity of each filter can be measured by its execution speed. This speed is directly related to the number of comparisons, additions, and multiplications required. It also depends, of course, on the algorithm structure, on the programming language, and on the computer characteristics. Therefore, the computational complexity of a filter depends on many factors and cannot be easily described by a single number.

Until now, no comprehensive comparison of all known order statistic filters has appeared in the literature. Most comparisons include at most 3-4 known nonlinear filters. Only a recent comparison includes a wide variety of nonlinear filters [236]. The results of the comparison are both quantitative and qualitative. Taking into account all the above-mentioned facts, we have decided to use the

following crude ranking of filter performance:

0: Poor performance

1: Average performance

2: Good performance.

For the bias, the ranking is the following:

0: Strong bias

1: Average bias

2: Low bias.

The following ranking has been used for computational complexity:

0: High computational complexity

1: Average computational complexity

2: Low computational complexity.

The filter comparison is summarized in Table 4. The entries to this table have been filled according to the characteristics of the filters given by the various researchers who have done comparisons (e.g., [236]) and according to the experience of the authors of this paper. This table is intended to be a qualitative guide for design and application engineers who want to choose a filter for a specific application fast and without much experimentation.

Certain subjects discussed in this review paper have already reached maturity, e.g. the analysis of the properties of the median filters. However, most of the subjects discussed are very active research areas. This fact explains the number of the new publications that appear in this area in international scientific journals and conferences. Nonlinear moving image processing is a rapidly expanding area. The development of digital HDTV will certainly give a big technological push in this area. Nonlinear multichannel signal/image processing has also a great potential in the near future. It is driven by the need for quality color image processing in the graphics industry as well as in digital video processing. It will also have several applications in multichannel signal processing (e.g., in geophysical applications, in multielectrode ECG/EEG processing and critical care monitoring). Adaptive nonlinear filtering is also very promising. It has a solid background from the theory of adaptive linear processing and a wide variety of applications in telecommunications (e.g. nonlinear echo cancellation, channel equalization), in medical signal processing, and in digital video filtering. From a theoretical point of view, the greatest challenge ahead is certainly the development of unifying theories for nonlinear filters. Such efforts have been made in the past (e.g. by using threshold decomposition and stack filters). However, the goal is difficult and such an achievement will be a breakthrough, with far reaching consequences in several research areas. Smaller steps toward this goal are more likely to appear in the near future by continuing current and past efforts. Merging of various nonlinear filter classes will give fruitful results in this direction. For example, a merging of order statistics filters with polynomial filters and neural networks can provide novel filter structures with enhanced adaptivity, good performance, and learning capabilities. The applications of order statistics filters are also increasing rapidly, ranging from digital TV image processing to sonar signal processing [237]–[241]. Computation speed is critical in most applications and poses new demands on fast chips specialized for order statistics filtering applications. Therefore, there is an ongoing effort to produce nonlinear filter implementations by using either analog or digital techniques. In general, order statistics image processing is a hot research and development area and will provide fruitful results and products in the near future.

REFERENCES

- [1] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ: Prentice-Hall, 1977.
- [2] M. D. Levine, *Vision in Man and Machine*. New York: McGraw-Hill, 1985.
- [3] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1978.
- [4] H. A. David, *Order Statistics*. New York: Wiley, 1981.
- [5] I. Pitas and A. N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*. Boston, MA: Kluwer Academic, 1990.
- [6] J. W. Tukey, *Exploratory Data Analysis*. Reading, MA: Addison-Wesley, 1970, 1977.
- [7] F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel, *Robust Statistics: An Approach Based on Influence Functions*. New York: Wiley, 1986.
- [8] P. S. Huber, *Robust Statistics*. New York: Wiley, 1981.
- [9] E. L. Lehmann, *Theory of Point Estimation*. New York: Wiley, 1983.
- [10] C. G. Boncelet, "Algorithms to compute order statistic distributions," *SIAM J. Stat. Comput.*, vol. 8, no. 5, pp. 868–876, Sept. 1987.
- [11] D. M. Raymond and M. M. Fahmy, "On the multivariate distributions of order statistics of discrete-state processes," in *Proc. 1991 IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Toronto), 1991, pp. 1717–1720.
- [12] D. M. Raymond and M. M. Fahmy, "Optimal M-D sorting using distributions of order statistics," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Toronto), 1991, pp. 2901–2904.
- [13] Y. H. Lee and S. Tantarata, "A simple formula for the variance of the median filter output with Laplacian input," *IEEE Trans. Signal Process.*, vol. 39, pp. 984–986, Apr. 1991.
- [14] P. Papantoni-Kazakos and R. M. Gray, "Robustness of estimators on stationary observations," *Ann. Prob.*, vol. 7, pp. 989–1002, 1979.
- [15] B. Kleiner, R. D. Martin, and D. J. Thomson, "Robust estimation of power spectra," *J. Roy. Statist. Soc. Ser. B*, vol. 41, pp. 313–351, 1979.
- [16] R. D. Martin and D. J. Thomson, "Robust resistant spectrum estimation," *Proc. IEEE*, vol. 70, pp. 1097–1115, Sept. 1982.
- [17] B. T. Poljak and Y. Z. Tsympkin, "Robust identification," *Automatica*, vol. 16, pp. 53–63, 1980.
- [18] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, pp. 433–481, Mar. 1985.
- [19] H. V. Poor, "On robust Wiener filtering," *IEEE Trans. Automat. Contr.*, vol. 25, pp. 531–536, June 1980.
- [20] K. S. Vastola and H. V. Poor, "Robust Wiener-Kolmogoroff theory," *IEEE Trans. Inform. Theory*, vol. 30, pp. 316–327, Mar. 1984.
- [21] S. A. Kassam, "Robust hypothesis testing and robust time series interpolation and regression," *J. Time Series Analysis*, vol. 3, pp. 185–194, 1982.
- [22] K. M. Ahmed and R. J. Evans, "Robust signal and array processing," *Proc. Inst. Elec. Eng.*, vol. 129, pt. F, no. 4, pp. 297–302, Aug. 1982.
- [23] A. H. El-Sawy and V. D. VandeLinde, "Robust detection of known signals," *IEEE Trans. Inform. Theory*, vol. 23, pp. 722–727, Nov. 1977.
- [24] J. W. Modestino, "Adaptive detection of signals in impulsive noise environments," *IEEE Trans. Commun.*, vol. 25, pp. 1022–1027, Sept. 1977.
- [25] B. I. Justusson, "Median filtering: Statistical properties," in *Two-Dimensional Digital Signal Processing II*, T. S. Huang, Ed. New York: Springer Verlag, 1981.
- [26] J. P. Fitch, E. J. Coyle, and N. C. Gallagher, "The analog median filter," *IEEE Trans. Circuits Syst.*, vol. 33, pp. 94–102, Jan. 1986.
- [27] H. G. Longbotham and A. C. Bovik, "Comments on the analog median filter," *IEEE Trans. Circuits Syst.*, vol. 36, p. 310, Feb. 1989.
- [28] F. Kuhlmann and G. L. Wise, "On the second moment properties of median filtered sequences of independent data," *IEEE Trans. Commun.*, vol. 29, pp. 1374–1379, Sept. 1981.
- [29] G. Y. Liao, T. A. Nodes, and N. C. Gallagher, "Output distributions of two-dimensional median filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, pp. 1280–1295, Oct. 1985.
- [30] T. A. Nodes and N. C. Gallagher, "The output distribution of the median type filters," *IEEE Trans. Commun.*, vol. 32, pp. 532–541, May 1984.
- [31] D. H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

- [32] G. J. Yong and T. S. Huang, "The effect of median filtering in edge location estimation," *Computer Vision, Graphics and Image Processing*, vol. 15, pp. 224-245, 1981.
- [33] E. Ataman, V. K. Aatre, and K. M. Wong, "Some statistical properties of median filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, pp. 1073-1075, Oct. 1981.
- [34] A. C. Bovik, T. S. Huang, and D. C. Munson, "The effect of median filtering on edge estimation and detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, pp. 181-194, Mar. 1987.
- [35] A. C. Bovik, "Streaking in median filtered images," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 493-503, Apr. 1987.
- [36] L. R. Rabiner, M. R. Sambur, and C. E. Schmidt, "Applications of a nonlinear smoothing algorithm to speech processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 23, pp. 552-557, Dec. 1975.
- [37] S. G. Tyan, "Median filtering: Deterministic properties," in *Two-Dimensional Signal Processing II*, T. S. Huang, Ed. New York: Springer Verlag, 1981.
- [38] G. R. Arce, N. C. Gallagher, and T. A. Nodes, "Median filters: Theory for one- and two-dimensional filters," in *Advances in Computer Vision and Image Processing*, T. S. Huang, Ed. JAI Press, 1986.
- [39] N. C. Gallagher and G. L. Wise, "A theoretical analysis of the properties of the median filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, pp. 1135-1141, Dec. 1981.
- [40] J. Astola, P. Heinonen, and Y. Neuvo, "On root structure of median and median-type filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 1199-1201, Aug. 1987.
- [41] A. Restrepo and A. C. Bovik, "Locally monotonic regression," in *Proc. 1989 IEEE Int. Conf. Acoustics, Speech and Signal Processing (Glasgow)*, 1989, pp. 1318-1321.
- [42] A. Restrepo and A. C. Bovik, "Non-Euclidean locally monotonic regression," in *Proc. 1990 IEEE Int. Conf. Acoustics, Speech and Signal Processing (Albuquerque)*, 1990, pp. 1201-1204.
- [43] G. Arce and N. C. Gallagher, "State description for the root-signal set of median filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 30, pp. 894-902, Dec. 1982.
- [44] G. R. Arce and N. C. Gallagher, "Root-signal set analysis for median filters," in *Proc. Allerton Conf. Commun. Contr. Comp.*, Oct. 1980.
- [45] J. P. Fitch, E. J. Coyle, and N. C. Gallagher, "Root properties and convergence rates for median filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, pp. 230-240, Feb. 1985.
- [46] D. Eberly, H. Longbotham, and J. Aragon, "Complete classification of roots to one-dimensional median and rank-order filters," *IEEE Trans. Signal Process.*, vol. 39, pp. 197-200, Jan. 1991.
- [47] P. D. Wendt, E. J. Coyle, and N. C. Gallagher, "Some convergence properties of median filters," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 276-286, Mar. 1987.
- [48] Z. J. Gan and M. Mao, "Two convergence theorems on the deterministic properties of median filters," *IEEE Trans. Signal Process.*, vol. 39, pp. 1689-1691, July 1991.
- [49] G. R. Arce and N. C. Gallagher, "BTC image coding using median filter roots," *IEEE Trans. Commun.*, vol. 31, June 1983.
- [50] J. P. Fitch, E. J. Coyle, and N. C. Gallagher, "Median filtering by threshold decomposition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 1183-1188, Dec. 1984.
- [51] G. R. Arce, "Statistical threshold decomposition for recursive and nonrecursive median filters," *IEEE Trans. Inform. Theory*, vol. IT-29, Mar. 1986.
- [52] G. R. Arce and R. L. Stevenson, "On the synthesis of median filter systems," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 420-429, Apr. 1987.
- [53] P. D. Wendt, E. J. Coyle, and N. C. Gallagher, Jr., "Stack filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, pp. 898-911, Aug. 1986.
- [54] E. J. Coyle, "Rank order operators and the mean absolute error criterion," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 63-76, Jan. 1988.
- [55] E. J. Coyle and J. H. Lin, "Stack filters and the mean absolute error criterion," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 1244-1254, Aug. 1988.
- [56] J. H. Lin and E. J. Coyle, "Minimum mean absolute error estimation over the class of generalized stack filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 663-678, Apr. 1990.
- [57] E. J. Coyle, J. H. Lin, and M. Gabbouj, "Optimal stack filtering and the estimation and structural approaches to image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 2037-2066, Dec. 1989.
- [58] M. Gabbouj and E. J. Coyle, "Minimum mean absolute error stack filtering with structural constraints and goals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 955-968, June 1990.
- [59] J. H. Lin, T. M. Sellke, and E. J. Coyle, "Adaptive stack filtering under the mean absolute error criterion," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 938-954, June 1990.
- [60] P. D. Wendt, "Nonrecursive and recursive stack filters and their filtering behavior," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 2099-2107, Dec. 1990.
- [61] P. M. Narendra, "A separable median filter for image noise smoothing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 3, pp. 20-29, Jan. 1981.
- [62] T. A. Nodes and N. C. Gallagher, "Median filters: Some modifications and their properties," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 30, pp. 739-746, Oct. 1982.
- [63] T. A. Nodes and N. C. Gallagher, "Two-dimensional root structures and convergence properties of the separable median filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, pp. 1350-1365, Dec. 1983.
- [64] G. Arce and N. C. Gallagher, "Stochastic analysis for the recursive median filter process," *IEEE Trans. Inform. Theory*, vol. 34, pp. 669-679, July 1988.
- [65] W. W. Boles, M. Kanewski, and M. Simaan, "Recursive two-dimensional median filtering algorithms for fast image root extraction," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 1323-1326, Oct. 1988.
- [66] R. J. Crinon and G. R. Arce, "Median filters: Analysis for signals with additive impulsive noise," in *Proc. 21st Allerton Conf.*, Oct. 1983.
- [67] M. P. Loughlin and G. R. Arce, "Deterministic properties of the recursive separable median filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 98-106, Jan. 1987.
- [68] G. R. Arce and R. J. Crinon, "Median filters: Analysis of two-dimensional recursively filtered signals," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 1984.
- [69] D. R. K. Brownrigg, "Weighted median filters," *Commun. Ass. Comput. Mach.*, vol. 27, pp. 807-818, Aug. 1984.
- [70] O. Yli-Harja, J. Astola, and Y. Neuvo, "Analysis of the properties of median and weighted median filters using threshold logic and stack filter representation," *IEEE Trans. Signal Process.*, vol. 39, pp. 395-410, Feb. 1991.
- [71] S. J. Ko and Y. H. Lee, "Center weighted median filters and their applications to image enhancement," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 984-993, Sept. 1991.
- [72] B. Zeng, Y. Neuvo, and A. N. Venetsanopoulos, "A class of weighted median filters for image processing," in *Proc. 1992 Conf. Information Sciences and Systems (Princeton, NJ)*, Mar. 18-20, 1992.
- [73] T. Loupas, W. N. McDicken, and P. L. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 129-136, Jan. 1989.
- [74] G. R. Arce and M. P. McLoughlin, "Theoretical analysis of the max/median filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 60-69, Jan. 1987.
- [75] A. Nieminen and Y. Neuvo, "Comments on the theoretical analysis of the max/median filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 826-827, May 1988.
- [76] G. R. Arce and R. E. Foster, "Detail preserving ranked-order based filters for image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 83-98, Jan. 1989.
- [77] G. Arce, "Multistage order statistics filters for image sequence processing," *IEEE Trans. Signal Process.*, vol. 39, pp. 1146-1163, May 1991.
- [78] M. B. Alp and Y. Neuvo, "3-dimensional median filters for image sequence processing," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (Toronto)*, 1991, pp. 2917-2920.
- [79] X. Wang and D. Wang, "On the max/median filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1473-1475, Aug. 1990.
- [80] P. Heinonen and Y. Neuvo, "FIR-median hybrid filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 832-838, June 1987.

- [81] J. Astola, P. Heinonen, and Y. Neuvo, "Linear median hybrid filters," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1430-1438, Nov. 1989.
- [82] P. Heinonen and Y. Neuvo, "FIR median hybrid filters with predictive FIR substructures," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 892-899, June 1988.
- [83] J. Neejarvi and Y. Neuvo, "Sinusoidal and pulse responses of the FIR-median hybrid filters," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 1552-1556, Dec. 1990.
- [84] A. Nieminen, P. Heinonen, and Y. Neuvo, "A new class of detail preserving filters for image processing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, pp. 74-90, Jan. 1987.
- [85] O. Vainio, Y. Neuvo, and S. E. Butner, "A signal processor for median-based algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 1406-1414, Sept. 1989.
- [86] R. Wichman, J. Astola, P. Heinonen, Y. Neuvo, "FIR median hybrid filters with excellent transient response in noisy conditions," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 2108-2117, Dec. 1990.
- [87] G. Heygster, "Rank filters in digital image processing," *Computer Vision, Graphics and Image Processing*, vol. 19, pp. 148-164, 1982.
- [88] J. Saniie, K. Donohue, and N. M. Bilgutay, "Order statistic filters as postdetection processors," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1722-1732, Oct. 1990.
- [89] J. P. Fitch, E. J. Coyle, and N. C. Gallagher, "Threshold decomposition of multidimensional ranked-order operations," *IEEE Trans. Circuits Syst.*, vol. 32, pp. 445-450, May 1985.
- [90] K. Chen, "Bit-serial realizations of a class of nonlinear filters based on positive Boolean functions," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 785-794, June 1989.
- [91] K. Chen, "Bit-serial realizations of a class of nonlinear filters using a bit-serial approach," in *Proc. 1988 IEEE Symposium on Circuits and Systems (Helsinki, Finland)*, pp. 1749-1752.
- [92] J. P. Fitch, "Software and VLSI algorithms for generalized ranked order filtering," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 553-559, May 1987.
- [93] V. Kim and L. Yaroslavskii, "Rank algorithms for picture processing," *Computer Vision graphics and Image Processing*, vol. 35, pp. 234-258, 1986.
- [94] J. Serra, *Image Analysis and Mathematical Morphology*. New York: Academic, 1982.
- [95] J. Serra, Ed., *Image Analysis and Mathematical Morphology: Theoretical Advances*, vol. 2. New York: Academic Press, 1988.
- [96] C. R. Giardina and E. R. Dougherty, *Morphological Methods in Image and Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1988.
- [97] E. R. Dougherty and C. R. Giardina, *Image Processing - Continuous to Discrete*, vol. 1. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [98] J. Serra, "Introduction to mathematical morphology," *Computer Vision, Graphics and Image Processing*, vol. 35, pp. 283-305, 1986.
- [99] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, pp. 532-550, July 1987.
- [100] P. Maragos, "Tutorial on advances in morphological image processing and analysis," *Opt. Eng.*, vol. 26, pp. 623-632, July 1987.
- [101] P. Maragos and R. W. Schafer, "Morphological systems for multidimensional signal processing," *Proc. IEEE*, vol. 78, pp. 690-710, Apr. 1990.
- [102] C. Lantuejoul and J. Serra, "M-filters," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 1982, pp. 2063-2066.
- [103] S. R. Sternberg, "Biological image processing," *Computer*, pp. 22-34, Jan. 1983.
- [104] S. R. Sternberg, "Grayscale morphology," *Computer Vision, Graphics and Image Processing*, vol. 35, pp. 333-355, 1986.
- [105] P. Maragos and R. W. Schafer, "Morphological filters, part I: Their set theoretic analysis and relations to linear shift invariant filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 1153-1169, Aug. 1987.
- [106] P. Maragos and R. W. Schafer, "Morphological filters, part II: Their relation to median, order statistics and stack filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 1170-1184, Aug. 1987.
- [107] R. L. Stevenson and G. R. Arce "Morphological filters: Statistic and further syntactic properties," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 1292-1305, Nov. 1987.
- [108] P. Maragos, "Pattern spectrum and multiscale shape representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 701-716, July 1989.
- [109] I. Pitas and A. N. Venetsanopoulos "Nonlinear mean filter in image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, pp. 573-584, June 1986.
- [110] J. B. Bednar and T. L. Watt "Alpha-trimmed means and their relationship to the median filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 145-153, Feb. 1987.
- [111] G. R. Arce and S. A. Fontana, "On the midrange estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 920-922, June 1988.
- [112] Y. H. Lee and S. A. Kassam "Generalized median filtering and related nonlinear filtering techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, pp. 672-683, June 1985.
- [113] Y. H. Lee and S. A. Kassam "Nonlinear edge preserving filtering techniques for image enhancement," in *Proc. 27th Midwest Symp. Circuits and Systems*, 1984, pp. 554-557.
- [114] A. C. Bovik, T. S. Huang, and D. C. Munson "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, pp. 1342-1349, Dec. 1983.
- [115] E. H. Loyd, "Least-squares estimation of location and scale parameters using order statistics," *Biometrika*, vol. 39, pp. 88-95, 1952.
- [116] L. Naaman and A. C. Bovik, "Least squares order statistic filters for signal restoration," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 244-257, Mar. 1991.
- [117] L. Naaman and A. C. Bovik, "Least squares order statistic filters for signal restoration in dependent noise," in *Proc. IEEE 1990 Int. Conf. Acoustics, Speech and Signal Processing (Toronto)*, 1991, pp. 1225-1228.
- [118] C. L. Mallows, "Resistant smoothing," in *Time Series*, O. D. Anderson, Ed. New York: North Holland, 1980, pp. 147-155.
- [119] C. L. Mallows, "Some theory of nonlinear smoothers," *Ann. Statist.*, vol. 8, no. 4, pp. 695-715, 1980.
- [120] L. Naaman and A. C. Bovik, "Optimal order statistic filters with coefficient censoring," in *Proc. ICASSP88*, 1988, pp. 820-823.
- [121] C. Kotropoulos and I. Pitas, "Nonlinear filtering of speckle noise in ultrasound B-mode images," in *Proc. Comp. Anal. Images and Patterns, CAIP91 (Dresden)*, 1991, pp. 118-123.
- [122] C. Kotropoulos and I. Pitas, "Optimum nonlinear detection and estimation in the presence of ultrasonic speckle," submitted to *Ultrasonic Imaging*, July 1992.
- [123] H. G. Longbotham and A. C. Bovik, "Theory of order statistic filters and their relationships to linear FIR filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 275-287, Feb. 1989.
- [124] H. G. Longbotham, A. C. Bovik, and A. Restrepo, "Generalized order statistic filters," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (Glasgow, Scotland)*, 1989, pp. 1610-1613.
- [125] F. Palmieri and C. G. Boncelet, "Ll-filters, A new class of order statistic filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 691-701, May 1989.
- [126] P. P. Gandhi and S. A. Kassam, "Nonlinear C-filters for nonstationary signals and deconvolution in impulsive noise," in *Proc. 1988 IEEE Symp. Circuits and Systems (Helsinki, Finland)*, pp. 1507-1510.
- [127] P. P. Gandhi, I. Song, and S. A. Kassam, "Nonlinear smoothing filters based on rank estimates of location," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 1359-1379, Sept. 1989.
- [128] R. J. Crinon, "The Wilcoxon filter: A robust filtering scheme," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Mar. 1985.
- [129] J. Song and S. A. Kassam, "Nonlinear filters based on generalized ranks for edge preserving smoothing," in *Proc. IEEE Symp. Circuits and Systems*, 1986, pp. 401-404.
- [130] P. P. Gandhi and S. A. Kassam, "Performance of some rank filters for edge preserving smoothing," in *Proc. IEEE Int. Symp. Circuits and Systems*, 1987, pp. 264-267.
- [131] P. P. Gandhi, I. Song, and S. A. Kassam, "Nonlinear smoothing filters based on rank estimates of location," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 1359-1379, Sept. 1989.
- [132] J. H. Lee and J. S. Kao, "A fast algorithm for two-dimensional Wilcoxon filtering," in *Proc. IEEE Symp. Circuits Syst.*, 1987,

- pp. 268–271.
- [133] A. Kundu and W. Wu, "Double window Hodges-Lehman (D) filter and hybrid D-median filter for robust image smoothing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 1293–1298, Aug. 1989.
- [134] A. Kundu and W. R. Wu, "Double window Hodges-Lehman (D) filter and hybrid D-median filter for robust image smoothing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 1293–1298, Aug. 1989.
- [135] S.R. Peterson, Y. H. Lee, and S. A. Kassam, "Some statistical properties of alpha-trimmed mean and standard type M -filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 707–713, May 1988.
- [136] C. A. Pomalaza-Raez and C. D. McGillen, "An adaptive nonlinear edge preserving filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 571–576, June 1984.
- [137] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
- [138] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [139] A. N. Oppenheim, R. W. Schaffer, and T. G. Stockham "Nonlinear filtering of multiplied and convolved signals," *Proc. IEEE*, vol. 56, pp. 1264–1294, Aug. 1969.
- [140] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. New York: Wiley, 1980.
- [141] G. Ramponi, "Quadratic filters for image enhancement," *Proc. Fourth European Signal Processing Conf., EUSIPCO-88* (Grenoble, France), 1988, pp. 239–242.
- [142] G. Ramponi, "Enhancement of low-contrast images by nonlinear operators," *Alta Frequenza*, vol. LVII, no. 7, pp. 451–455, Sept. 1988.
- [143] G. Ramponi and G. Sicuranza, "Quadratic digital filters for image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 937–939, June 1988.
- [144] G. Ramponi, "Edge extraction by a class of second-order nonlinear filters," *Electron. Lett.*, vol. 22, no. 9, Apr. 24, 1986.
- [145] G. Ramponi and G. Sicuranza, "Decision-directed nonlinear filter for image processing," *Electron. Lett.*, vol. 23, no. 23, pp. 1218–1219, Nov. 5, 1987.
- [146] G. Ramboni, "Bi-impulse response design of isotropic quadratic filters," *Proc. IEEE*, vol. 78, no. 4, pp. 665–677.
- [147] A. Kundu, S. K. Mitra, and P. P. Vaidyanathan, "Application of two-dimensional generalized mean filtering for the removal of impulse noises from images," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 600–608, June 1984.
- [148] I. Pitas and A. N. Venetsanopoulos, "A new filter structure for the implementation of certain classes of image operations," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 636–647, June 1988.
- [149] G. Angelopoulos and I. Pitas, "Multichannel Wiener filters in color image restoration based on AR modeling," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing* (Albuquerque, NM), 1990, pp. 2517–2520.
- [150] I. Pitas and A. N. Venetsanopoulos "Nonlinear order statistic filter for image filtering and edge detection," *Signal Processing*, vol. 10, pp. 395–413, June 1986.
- [151] B. R. Hunt, "Karhunen-Loeve multispectral image restoration, part I: Theory," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 592–599, June 1984.
- [152] N. P. Galatsanos and R. T. Chin, "Digital restoration of multichannel images," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 415–421, Mar. 1989.
- [153] G. Aggelopoulos and I. Pitas, "Least-squares multichannel filters in color image restoration," in *Proc. European Conf. Circuit Theory and Design ECCTD89* (Brighton, England), Sept. 1989.
- [154] C. W. Therrien, "Multichannel filtering methods for color image segmentation," in *Proc. 1985 IEEE Conf. Computer Vision and Pattern Recognition*, pp. 637–639.
- [155] N. Ohyama, M. Yachida, E. Badique, J. Tsujiuchi, and T. Honda, "Least squares filter for color image segmentation," *J. Opt. Soc. Amer. A*, vol. 5, no. 1, pp. 19–24, Jan. 1988.
- [156] I. Pitas, "Marginal order statistics in color image processing," *Opt. Eng.*, vol. 29, no. 5, pp. 495–503, May 1990.
- [157] I. Pitas, "Marginal order statistics in color image processing," in *Proc. European Conf. Signal Processing* (Barcelona), 1990, pp. 809–812.
- [158] I. Pitas and P. Tsakalides, "Multivariate ordering in color image restoration," *IEEE Trans. Circuits Syst. for Video Technology*, vol. 1, pp. 247–259, Sept. 1991.
- [159] C. Kotropoulos and I. Pitas, "Multichannel L-filters based on marginal data ordering," submitted to *IEEE Trans. Signal Process.*
- [160] J. Astola, P. Haavisto, P. Heinonen, and Y. Neuvo, "Median type filters for color signals," in *Proc. IEEE Int. Symp. Circuits and Systems*, 1988, pp. 1753–1756.
- [161] K. Oistamo, P. Jarske, J. Astola, and Y. Neuvo, "Vector median filters for complex signals," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing* (Glasgow), 1989, pp. 813–816.
- [162] J. Astola, P. Haavisto, and Y. Neuvo, "Vector median filters," *Proc. IEEE*, vol. 78, pp. 678–690, Apr. 1990.
- [163] S. Sanwalka and A. N. Venetsanopoulos, "Vector order statistic filtering of colour images," in *Proc. Thirteenth GRETSI Symp. Signal and Image Processing* (Juan-les Pins, France), Sept. 16–20, 1991.
- [164] V. Barnett, "The ordering of multivariate data," *J. Roy. Statist. Soc. A*, vol. 139, pt. 3, pp. 318–354, 1976.
- [165] A. M. Mood, "On the joint distribution of the medians in samples from a multivariate population," *Ann. Math. Statist.*, vol. 12, pp. 268–279, 1941.
- [166] C. M. Mustafi, "A recurrence relation for distribution functions of order statistics from bivariate distributions," *J. Amer. Statist. Ass.*, vol. 64, pp. 600–601, 1969.
- [167] J. Galambos, "Order statistics of samples from multivariate distributions," *J. Amer. Statist. Ass.*, vol. 70, pp. 674–680, 1975.
- [168] R. C. Hardie and G. R. Arce, "Ranking in R^p and its use in multivariate image estimation," *IEEE Trans. Circuits Syst. for Video Technology*, vol. 1, pp. 197–209, June 1991.
- [169] F. P. Preparata and M. I. Shamos, *Computational Geometry*. New York: Springer Verlag, 1986.
- [170] G. A. F. Seber, *Multivariate Observations*. New York: Wiley, 1984.
- [171] J. S. Lee, "Digital image enhancement and noise filtering by local statistics," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 2, pp. 165–168, Mar. 1980.
- [172] J. S. Lee, "Refined filtering of image noise using local statistics," *Computer Graphics and Image Processing*, vol. 15, pp. 380–389, 1981.
- [173] X. Z. Sun and A. N. Venetsanopoulos, "Adaptive schemes for noise filtering and edge detection by use of local statistics," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 57–69, Jan. 1988.
- [174] I. Scollar, B. Weidner, and T. S. Huang, "Image filtering using the interquartile distance," *Computer Vision, Graphics and Image processing*, vol. 25, pp. 236–251, 1984.
- [175] R. Ding and A. N. Venetsanopoulos, "Generalized homomorphic and adaptive order statistic filters for the removal of impulsive and signal dependent noise," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 948–955, Aug. 1987.
- [176] I. Pitas and A. N. Venetsanopoulos, "Edge detectors based on order statistics," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 538–550, July 1986.
- [177] A. C. Bovik and D. C. Munson, "Edge detection using median comparisons," *Computer Vision Graphics and Image Processing*, vol. 33, pp. 377–389, 1986.
- [178] J. Hajek and Z. Sidak, *Theory of Rank Tests*. New York: Academic Press, 1967.
- [179] A. C. Bovik, T. S. Huang, and D. C. Munson, "Nonparametric tests for edge detection in noise," *Pattern Recognition*, vol. 19, no. 3, pp. 209–219, 1986.
- [180] R. Martens and A. N. Venetsanopoulos, "A comparison of adaptive nonlinear filters using different images," in *Proc. IEEE Int. Conf. Systems Engineering* (Dayton, OH), 1989.
- [181] A. Restrepo and A. C. Bovik, "Adaptive trimmed mean filters for image restoration," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 8, pp. 1326–1337, 1988.
- [182] R. V. Hogg, "More light on the kurtosis and related statistics," *J. Amer. Statist. Assoc.*, vol. 67, pp. 422–424, 1972.
- [183] Y. H. Lee and S. Tantarana, "Decision-based order statistic filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 406–420, Mar. 1990.
- [184] F. Cheng and A. N. Venetsanopoulos, "Adaptive nonlinear operators and image processing based on image geometrical structures," in *Proc. Int. Symp. Signal, Systems and Electronics '92* (Paris, France), Sept. 1992.
- [185] F. Cheng and A. N. Venetsanopoulos, "An adaptive morphological filter for image processing," *IEEE Trans. Image Processing*, to be published.

- [186] C. Rey and R. K. Ward, "A parametrized family of nonlinear image smoothing filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 1458-1462, Sept. 1989.
- [187] S. Y. Park and Y. H. Lee, "Double smoothing of images using median and Wiener filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 943-946, June 1989.
- [188] R. Bernstein, "Adaptive nonlinear filters for simultaneous removal of different kinds of noise in images," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 1275-1291, Nov. 1987.
- [189] R. Martens and A. N. Venetsanopoulos, "Adaptive order statistic filters: The complexity/quality trade off," presented at Midwest CAS Symposium, Naval Postgraduate School, Monterey, CA, May 14-17, 1991.
- [190] T. Alexander, *Adaptive Signal Processing*. New York: Springer Verlag, 1986.
- [191] M. Bellanger, *Adaptive Digital Filters and Signal Analysis*. New York: Marcel Dekker, 1987.
- [192] I. Pitas and A. N. Venetsanopoulos, "Adaptive filters based on order statistics," *IEEE Trans. Signal Process.*, vol. 39, pp. 518-522, Feb. 1991.
- [193] F. Palmieri and C. G. Boncelet, Jr., "A class of nonlinear adaptive filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (New York), 1988, pp. 1483-1486.
- [194] I. Pitas and A. N. Venetsanopoulos, "Adaptive L-filters," in *Proc. European Conf. Circuit Theory and Design* (Brighton, England), 1989, pp. 580-584.
- [195] I. Pitas and A. N. Venetsanopoulos, "LMS and RLS adaptive L-filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Albuquerque), 1990, pp. 1389-1392.
- [196] C. Kotropoulos and I. Pitas, "Constrained adaptive LMS L-filters," *Signal Processing*, vol. 26, pp. 335-358, 1992.
- [197] C. Kotropoulos and I. Pitas, "Design of the constrained adaptive LMS L-filters," in *Proc. MELECON* (Ljubljana), 1991.
- [198] C. Kotropoulos and I. Pitas, "Constrained adaptive LMS L-filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Toronto), 1991, pp. 1665-1668.
- [199] D. E. Rumelhardt and J. L. McClelland, Eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press, 1986.
- [200] F. Palmieri "A backpropagation algorithm for multilayer hybrid order statistics filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Glasgow, Scotland), 1989.
- [201] I. Pitas and S. Vougioukas, "Adaptive nonlinear filters based on order statistics," in *Proc. European Signal Processing Conf.* (Barcelona), 1990, pp. 397-400.
- [202] I. Pitas and S. Vougioukas, "LMS order statistics filter adaptation by backpropagation," *Signal Processing*, vol. 25, pp. 319-335, 1991.
- [203] F. Palmieri, "Adaptive recursive order statistics filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Albuquerque), 1990, pp. 1229-1232.
- [204] F. Palmieri and R. E. Croteau, "Adaptive channel equalization using generalized order statistic filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Toronto), 1991, pp. 1933-1936.
- [205] T. I. Haweel and P. M. Clarkson, "Analysis and generalization of a median adaptive filter," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Albuquerque), 1990, pp. 1269-1272.
- [206] G. A. Williamson and P. M. Clarkson, "On the convergence and steady state behavior in the median LMS adaptive filter," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Toronto), 1991, pp. 2149-2152.
- [207] D. E. Knuth, *The Art of Computer Programming*, vol. 3. Reading, MA: Addison-Wesley, 1973.
- [208] K. Ofazer, "Design and implementation of a single chip 1-D median filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, pp. 1164-1168, Oct. 1983.
- [209] J. Alsford et al., "CRS image processing system with VLSI modules," in *Image Processing System Architectures*, J. Kittler and M. Duff, Eds. Research Studies Press, 1985.
- [210] D. S. Richards, "VLSI median filters," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 38, pp. 145-153, Jan. 1990.
- [211] P. E. Allen, *Switched Capacitor Circuits*. New York: Van Nostrand Reinhold, 1984.
- [212] D. F. Franke, Ed., *Charge Coupled Devices*. New York: Springer Verlag, 1980.
- [213] D. R. Morgan, "Analog sorting networks ranks inputs by amplitudes and allows selection," *Electronic Design* vol. 2, pp. 72-74, Jan. 1973.
- [214] T. Jarske, P. Heinonen, and Y. Neuvo, "Switched-capacitor linear median hybrid filters," in *Proc. IEEE Int. Symp. Circuits and Systems*, 1987, pp. 260-263.
- [215] J. S. J. Ly and W. H. Holmes, "Analog implementation of median filters," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 1032-1033, Aug. 1988.
- [216] O. Vaino, Y. Neuvo, and S. E. Butner, "A signal processor for median-based algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 1406-1414, Sept. 1989.
- [217] K. Chen, P. Heinonen, Q. Ye, and Y. Neuvo, "Analog/digital hybrid median filter realizations," in *Proc. 1987 Int. Conf. Digital Signal Processing*, (Florence, Italy), 1987, pp. 349-352.
- [218] E. Horowitz and S. Sahni, *Fundamentals of Computer Algorithms*. Rockville, MD: Computer Science Press, 1978.
- [219] T. S. Huang, G. J. Yang, and G. Y. Tang, "A fast two-dimensional median filtering algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 27, pp. 13-18, Feb. 1979.
- [220] M. O. Ahmad and D. Sundararajan, "A fast algorithm for two-dimensional median filtering," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 1364-1374, Nov. 1987.
- [221] I. Pitas "Fast algorithms for running ordering and max/min calculation," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 795-804, June 1989.
- [222] S. J. Ko, Y. H. Lee, and A. T. Fam, "Efficient implementation of one-dimensional recursive median filters," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 1447-1450, Nov. 1990.
- [223] S. Ranka and S. Sahni, "Efficient serial and parallel algorithms for median filtering," *IEEE Trans. Signal Process.*, vol. 39, pp. 1462-1466, June 1991.
- [224] F. Pasian, "Sorting algorithms for filters based on ordered statistics: performance considerations," *Signal Processing*, vol. 14, pp. 287-293, 1988.
- [225] M. Juhola, J. Katajainen, and T. Raita, "Comparison of algorithms for standard median filtering," *IEEE Trans. Signal Process.*, vol. 39, pp. 204-208, Jan. 1991.
- [226] P. E. Danielsson "Getting the median faster," *Computer Graphics and Image Processing*, vol. 17, pp. 71-78, 1981.
- [227] E. Ataman, V. K. Aatre, and K. M. Wong "A fast method for real time median filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, pp. 415-421, Aug. 1980
- [228] R. T. Hoctor and S. A. Kassam, "An algorithm for order statistic determination and L-filtering," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 1988, pp. 1686-1689.
- [229] O. Yli-Harja, J. Astola, and Y. Neuvo, "Generalization of the radix-method of finding the median to weighted median, order statistic and weighted order statistic filters," *SPIE Visual Communication and Image Processing*, vol. 1001, pp. 69-75, 1988.
- [230] V. V. B. Rao and K. S. Rao, "A new algorithm for real-time median filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, pp. 1674-1675, Dec. 1986.
- [231] R. T. Hoctor and S. A. Kassam, "An algorithm and a pipelined architecture for order statistic determination and L-filtering," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 344-352, Mar. 1989.
- [232] C. F. Hall, "Subjective evaluation of a perceptual quality metric," *Image Quality, Proc. Soc. Photo Opt. Instr. Eng.*, vol. 310, pp. 200-204, 1981.
- [233] H. Marmolin, "Subjective MSE measures," *IEEE Trans. Syst., Man, Cybern.*, vol. 16, no. 3, pp. 486-489, 1986.
- [234] C. F. Hall, "The application of human visual system models to digital color image compression," in *Proc. IEEE Int. Commun. Conf.* (Boston), 1983, pp. 436-441.
- [235] L. Akarun and R. A. Haddad, "Performance measures for the rank order filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Toronto), 1991, pp. 2537-2540.
- [236] Y. S. Fong, C. A. Pomalaza, and X. H. Wang, "Comparison study of nonlinear filters in image processing applications," *Opt. Eng.*, vol. 28, no. 7, pp. 749-760, July 1989.
- [237] S. S. Perlman, S. Eisenhandler, P. W. Lyons, and M. J. Shumila, "Adaptive median filtering for impulse noise elimination in real-time TV signals," *IEEE Trans. Commun.* vol. 35, pp. 646-652, June 1987.
- [238] K. J. Hahn and K. M. Wong, "Median filtering of cepstra," in *Proc. 1983 Int. Electrical and Electronics Conf.*, (Toronto, Canada), 1983, pp. 352-355.
- [239] P. Zamperoni, "Feature extraction by rank-order filtering for image segmentation," *Int. J. Pattern Analysis and Artificial*

- Intelligence*, vol. 2, no. 2, pp. 301–319, 1988.
- [240] S. S. H. Naqvi, N. C. Gallagher, and E. J. Coyle, "An application of median filter to digital TV," in *Proc. 1986 IEEE Int. Conf. Acoustics, Speech and Signal Processing* (Tokyo, Japan), 1986.
- [241] K. M. Wong, S. Chen, "Detection of narrow-band sonar signals using order statistic filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 597–613, May 1987.

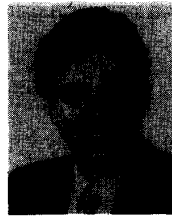


Ioannis Pitas received the diploma of electrical engineering in 1980 and the Ph.D. degree in electrical engineering in 1985, both from the University of Thessaloniki, Greece.

Since 1989 he has been an Assistant Professor at the University of Thessaloniki. From 1980 to 1989 he was scientific assistant and lecturer there. He served as a Visiting Research Associate at the University of Toronto, Canada, and at the University of Erlangen-Nürnberg, Germany, and was a Visiting Assistant Professor at the

University of Toronto. He was lecturer in short courses for continuing education. His current interests are in the areas of digital image processing, multidimensional signal processing, and computer vision.

Dr. Pitas has published over 100 papers and has contributed to five books. He is coauthor of the book *Nonlinear Digital Filters: Principles and Applications* (Kluwer Academic, 1990) and editor of the book *Parallel Algorithms and Architectures for Digital Image Processing, Computer Vision and Neural Networks* (J. Wiley, 1992, in press). He has served as a member of the European Community ESPRIT Parallel Action Committee and has been a member of the program committees of several scientific conferences and workshops.



Anastasios N. Venetsanopoulos (Fellow, IEEE) received the B.S. degree from the National Technical University of Athens (NTU), Greece, in 1965 and the M.S., M.Phil., and Ph.D. degrees in electrical engineering, all from Yale University, in 1966, 1968, and 1969, respectively.

He joined the University of Toronto, Canada, in September 1968 and is now a Professor in the Department of Electrical Engineering there. He also served as Chairman of the Communications Group (1974–1978 and 1981–1986)

and as Associate Chairman of the Department of Electrical Engineering (1978–1979). He was on research leave at the Swiss Federal Institute of Technology, the University of Florence, the Federal University of Rio de Janeiro, the National Technical University of Athens, and the Imperial College of Science and Technology and was an Adjunct Professor at Concordia University. He served as Lecturer of numerous short courses to industry and continuing education programs. Dr. Venetsanopoulos has contributed to 11 books and has published more than 300 papers in digital signal and image processing and digital communications. He has served as consultant to several organizations and was Editor of the *Canadian Electrical Engineering Journal* from 1981 to 1983.

Dr. Venetsanopoulos was President of the Canadian Society of Electrical Engineering and Vice President of the Engineering Institute of Canada (1983–1984). He has been a Fulbright Scholar, an A. F. Schmitt Scholar, and a recipient of the J. Vakis Award. He is a member of the New York Academy of Sciences, Sigma Xi, the International Society for Optical Engineering, and the Technical Chamber of Greece. Dr. Venetsanopoulos is a registered professional engineering in Ontario and Greece and a fellow of the Engineering Institute of Canada.